Optimization of poetry generation neural networks

Rick Roslof
WSU ID: 11810429
Email: richard.roslof@wsu.edu

Abstract

Poetry generation is a frequently studied problem in natural language processing. Several neural network-based approaches have been developed for this task, including Recurrent, Long Short-Term Memory, Gated Recurrent Units, and Transformer models. While there are an abundance of projects using these tools, they fail to provide uniformity or guidance on architectural specifics and parameter optimization. The aim of this project is to build a basic poetry generation network that can be iterated programmatically to determine the most efficient model architectures and parameter values.

Introduction

Text analysis and generation are unique challenges for artificial systems due to the need to preserve semantic context throughout a statement. Neural networks offer several sequential data processing architectures that excel at these tasks, including Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), and Transformer models. Each of these models provides a distinct approach to carrying data forward to each subsequent step of the training process.

RNNs differ from other feedforward networks by including a basic feedback loop within their network layer, allowing context to be maintained for the next calculation. However, they struggle with longer-term context and gradient vanishing problems. LSTM networks address these weaknesses by providing more sophisticated mechanisms for memory management, improving accuracy at the cost of requiring more parameter training than RNNs. GRUs serve as a middle-ground between the two, achieving most of the performance of LSTMs with a simpler memory system that reduces training times. Transformer models are much more complex, providing the training parallelization that underlies models such as Chat-GPT.[1]

While examples of such networks are easy to find, architectural design for poetry generation projects is more of an art than a science. The counts, types, and sizes of the network layers appear to be chosen based on the creator's intuition, with any iteration history omitted from the final reports. Additionally, most authors' suggestions for future improvement are that "more is better": more data, more epochs, more parameters.

This project seeks to improve the efficiency of existing poetry generation strategies, identifying the optimal balance between output quality and training time. After first creating a baseline generation network, an automated tool will be used to vary the network characteristics and record the resulting output. These results will be compared and presented in the final report.

Literature Review

A survey of existing projects reveals the large variety of architectural designs. These differences are summarized in Appendix A and include network type, depth, character vs. word generation, and other features such as dropout and bidirectionality.

Yu[2] analyzed the relative performance of RNN, LSTM, and GRU networks on poetry generation in multiple languages, showing that bi-directional networks of all types had superior training scores but produced worse text than single-directional networks, with attention mechanisms negatively impacting all results.

Kumar[3] also compared word-generation performance on deep RNN, LSTM, GRU, and Transformer networks, but did not explore further architectural alterations.

Trekhleb[4] showed excellent results with a shallow LSTN network after extensive preprocessing of the input data.

Tam[5] included helpful instructions for character-based vectorization of input text into pytorch-compatible training tensors.

Belli[6] created a parameterized recurrent language model in pytorch, which was a useful reference.

Morgan[7] served as a source of information about using perplexity as a metric for measuring generative language model quality.

Technical Plan

- **Hardware:** The training hardware available for this project was an AMD Ryzen 5 7600X CPU.
- **Data Sources:** Project Gutenberg's corpus of English-language poetry[8].
- **Libraries:** pytorch was used for network construction.
- **Preprocessing:** Cleaning, vectorization, and sequence batching were performed for both character- and word-based generation strategies. As the chosen corpus contains over one hundred million characters, options were provided to train on a subset of the data.
- **Model Creation:** A parameterized model was created to facilitate easy comparisons between network architectures. This model allowed for simple specification of layer sizes, recurrence layer counts & bidirectionality, dropout after each recurrence layer, and number of nonlinear fully-connected layers.
- **Training:** All models were trained using cross-entropy loss with the Adam optimizer. Training times and perplexity were recorded per-epoch to measure the impact of architecture changes.

- **Text Generation:** Output text was generated for each model from the following three input seeds:
    - "Two roads diverged in a yellow wood"
    - "And on the pedestal these words appear"
    - "Shall I compare thee to a summer's day?"
- **Iteration:** Many model variations were trained to compare the impacts of each architectural parameter. Each combination of recurrence type (RNN, LSTM, GRU) and token type (character, word) was trained using nine sets of parameters:
    - Baseline: One recurrence layer, no bidirectionality or dropout, no nonlinear layers.
    - Two recurrence layers.
    - Three recurrence layers.
    - 20% dropout.
    - 50% dropout.
    - Bidirectional.
    - One nonlinear layer.
    - Two nonlinear layers.
    - Two recurrence layers, 20% dropout, bidirectional, one nonlinear layer.

  Character-based models were trained on a random block of 100k characters read from the data set, while word-based models—which result in fewer tokens-per-character—were tokenized from a block of 200k characters to produce similar training times.

Complete Results

Results

Figures 1 through 3 show the relative perplexity values obtained versus training time for character-based models. Figures 4 through 6 show the same for word-based models, except using logarithmic perplexity due to the larger vocabulary sizes. Appendix B contains the raw data and architecture parameters for each model.

Appendix C contains each model's output text for each of the three input seeds. General trends in the output results are discussed below.

Architecture Impact Analysis

Recurrence type

As expected, RNN models were fastest to train but produced the lowest quality output and were more susceptible to gradient vanishing issues. However, the training time benefit of GRU over LSTM models was not as substantial as anticipated. Considering that LSTM models tended to produce slightly higher-quality results, the theoretical efficiency benefits of GRU models were not borne out in this testing.

Baseline models of all three recurrence types were prone to gradient vanishing.

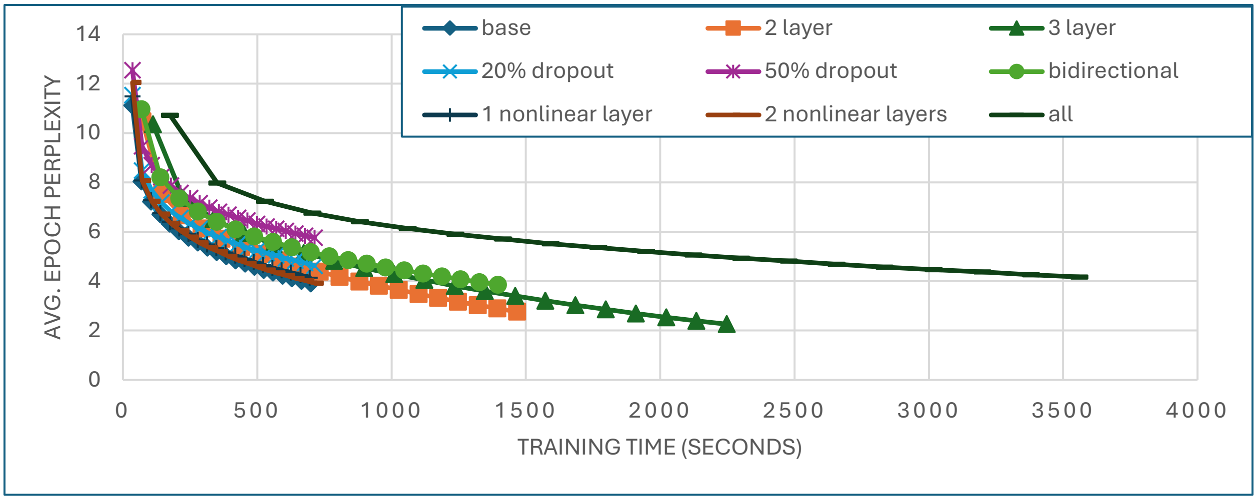

Figure 1: Character-based RNN model training data.

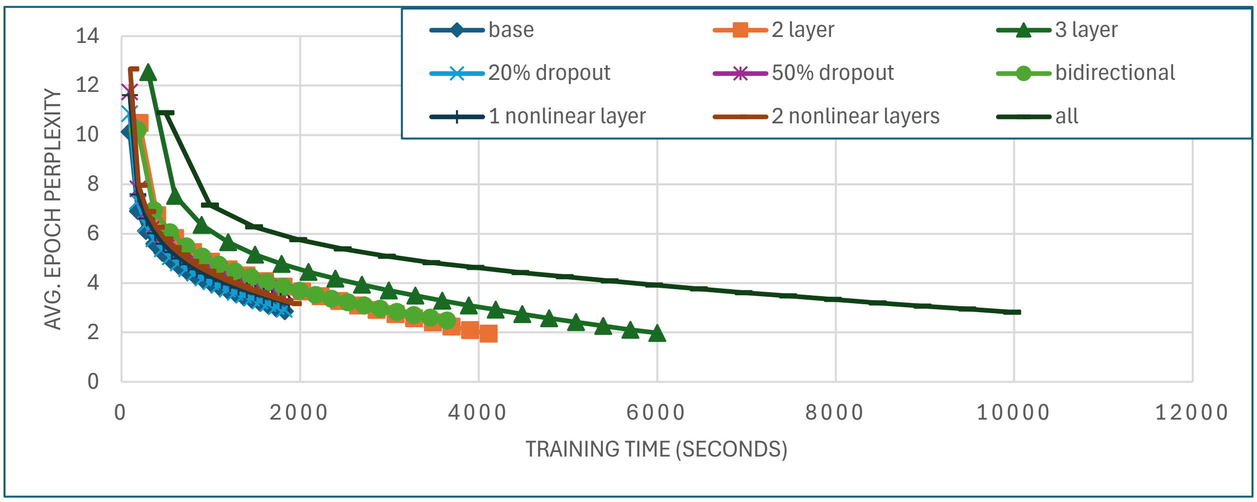

Figure 2: Character-based LSTM model training data.

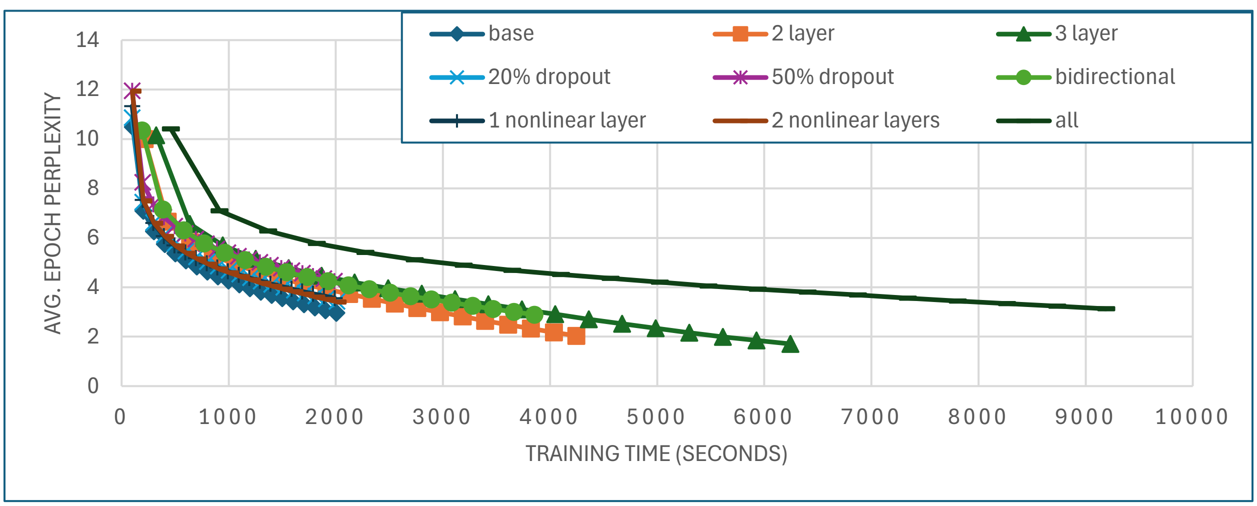

Figure 3: Character-based GRU model training data.

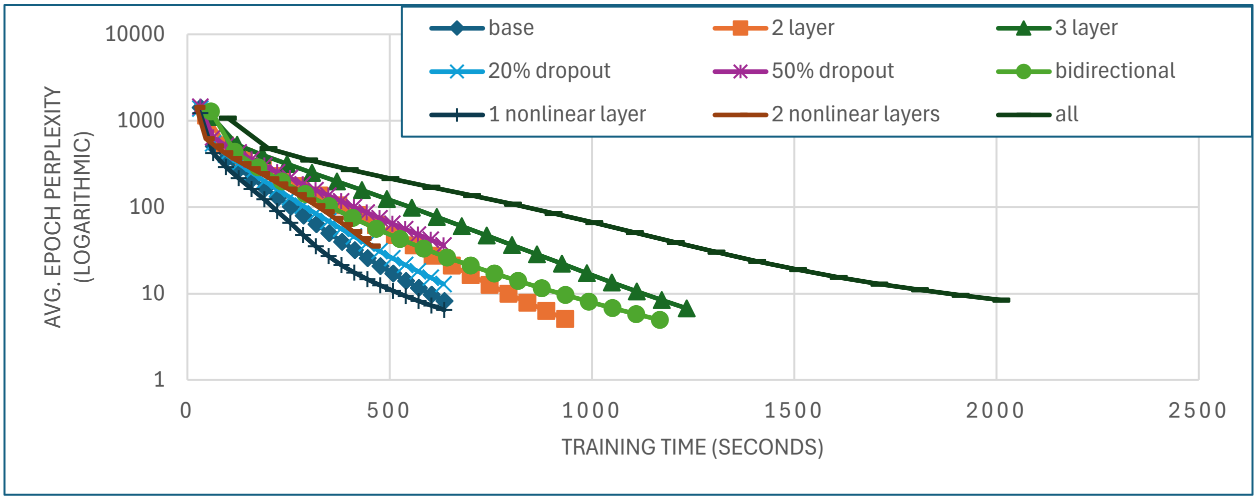

Figure 4: Word-based RNN model training data.

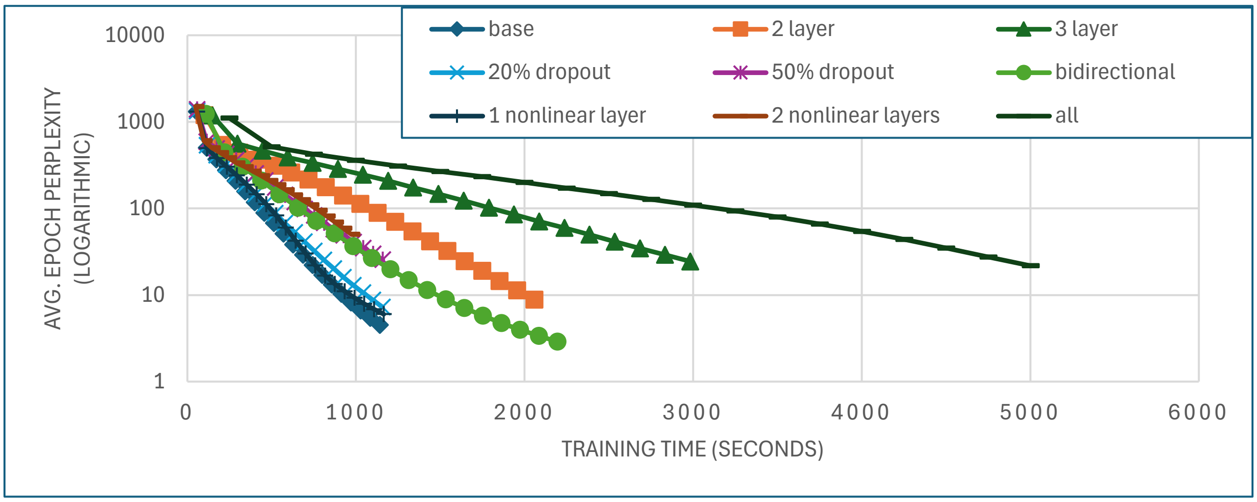

Figure 5: Word-based LSTM model training data.

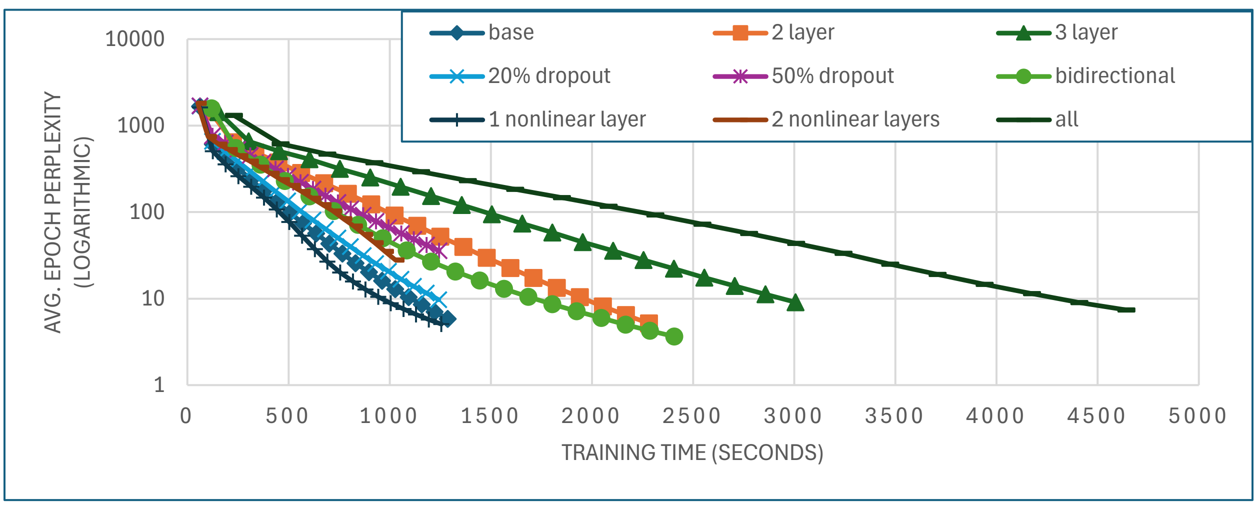

Figure 6: Word-based GRU model training data.

Token type

Character-based models were significantly more likely to experience gradient vanishing issues in their output text, especially using the baseline architecture parameters. Word-based models also demonstrated a higher baseline quality as they were inherently free of spelling errors, but the use of additional architecture parameters shrunk that gap to the point where both character- and word-based LSTM models were able to produce plausible poetry.

Recurrence layer count

As expected, additional layers increased training time proportionally to the expanded feature size. Character-based RNN and LSTM models still experienced gradient vanishing, but GRU models had their outputs significantly bolstered by a second recurrence layer. However, adding a third layer did not further improve output quality.

Dropout

Dropout was the single most impactful architecture parameter tested. While it introduced more spelling errors in character-based models, even 20% dropout was sufficient to eliminate gradient vanishing while only marginally worsening perplexity. Increasing dropout to 50% did not result in any further improvements.

Bidirectionality

Not many prior work used bidirectionality for poetry generation, evidently for good reason. The increase in trainable parameters tended to cause gradient problems even in word-based models. More investigation may be warranted into the slight improvement that word-based LSTM models experienced, but overall the architecture choice lengthened training times to no benefit.

Nonlinear layers

As with bidirectionality, the addition of nonlinear layers caused gradient issues in nearly all models. While training time was largely unaffected, only the "all parameters" models were still able to achieve good performance.

Limitations

The primary weakness of this analysis stems from the project's time constraints. Comparing a large number of architectures required training an equivalent number of models, which limited the amount of time it was possible to spend on each model's training. While distinct training trends are observable in the gathered data, many models were not trained to convergence and therefore the theoretical maximum quality obtainable by each hyperparameter set is unknown. A higher learning rate may have been able to address this issue at the risk of greater instability, but once the problem was noticed there was insufficient time to rerun all the previous training.

Secondarily, the difficulty in objectively assessing poetry quality (beyond determining if gradient vanishing occurred) limits the ability to assess output strength as a product of training time. Perplexity is a decent metric, but must be used in conjunction with subjective analysis to determine what is actually "good poetry".

Future Work After This Course

Additional insights can be obtained using the parameterized model created for this project. The application supports user-specified layer sizes, which would have been included in the data within this report if time had permitted. More testing will be performed by varying these values, with a focus on how those changes interact with the other architectural parameters. It is possible, for example, that reducing the recurrence layer feature size by half while doubling the layer count could produce better results within the same training time. Likewise, there may be some benefits of tuning layer sizes relative to the input vocabulary for word-based models; since word tokenization produces large vocabulary sizes, it is possible that the layer sizes used in previous testing were too small and thus constrained the model's performance.

It will also be simple to add features and hyperparameters that were not included in the original model. Gradient clipping has the best chance of improving model outputs, potentially preventing the output looping seen in the simpler model variations and is available via a pytorch function for easy implementation. Allowing the user to select a learning rate rather than using a single hard-coded value is another obvious next step to improve the application design.

Future Work

There are several avenues for other researchers to expand on the results presented here. While one of the goals for pursuing efficiency was to maximize the output quality of my personal hardware, it would still be within the spirit of the project to train models on a consumer GPU. This would speed up training times dramatically, allowing more permutations to be compared and better conclusions reached.

Another approach would be to improve the parameterized model. While it worked well for queuing the training of multiple parameter sets, the existing model has distinct limitations in the resulting network configurations. Increasing the network complexity without overwhelming the user with choices is not an easy task, but it could answer some interesting questions. One such open question is how combining different recurrence layer types affects the network's learning ability; since RNNs are fast but poor at long-term recall, perhaps being pairing one with a LSTM layer would allow for the best of both architectures.

Finally, this project could be extended to include Transformer models. LLMs are notorious for their power and training time requirements, so any efficiency optimizations could have significant impacts.

References

[1] GeeksForGeeks. “RNN vs LSTM vs GRU vs Transformers.” 2025. https://www.geeksforgeeks.org/rnn-vs-lstm-vs-gru-vs-transformers/

[2] Yu, Fengcheng. “From RNN to GRU and LSTM: A Study on the Performance of Poem Generation Models Integrating Bidirectional and Attention Techniques.” 2024. https://github.com/ffengc/NLP-PoetryModels

[3] Kumar, Ayush. “Text Generation Using RNN, GRU, LSTM, and Transformer Models.” 2025. https://github.com/ay05h/Text-Generation-

[4] Trekhleb, Oleksii. “Interactive Machine Learning Experiments.” https://github.com/trekhleb/machine-learning-experiments

[5] Tam, Adrian. “Text Generation with LSTM in PyTorch.” 2023. https://machinelearningmastery.com/text-generation-with-lstm-in-pytorch/

[6] Belli, Davide. “Word-level language modeling RNN.” 2017. https://github.com/davide-belli/nlp-RNN-pytorch

[7] Morgan, Abby. “Perplexity for LLM Evaluation.” 2024. https://www.comet.com/site/blog/perplexity-for-llm-evaluation/

[8] Parrish, Allison. “A Gutenberg Poetry Corpus.” 2019. https://github.com/aparrish/gutenberg-poetry-corpus

[9] Shah, Hrishita. “DeepPoet – An AI Poetic Text Generator using RNNs in Python.” 2023. https://github.com/Hrishita-Shah/DeepPoet---An-AI-Poetic-Text-Generator-using-RNNs-in-Python

[10] DeJans Jr., Adam. “LSTM Shakespeare-like Text Generation.” 2020. https://github.com/addejans/LSTM-Shakespeare-like-Text-Generation

[11] Abou-Zamzam, Zafir. “ShakespeareSonnetGeneration.” 2018. https://github.com/reislerjul/ShakespeareSonnetGeneration

[12] Costolo, Erin. “Shakespeare Sonnet Generator.” 2022. https://github.com/ErinNC/Shakespeare_Sonnet_Generator

[13] Brundyn, Annika. “The Beginner’s Guide to Recurrent Neural Networks and Text Generation.” 2018. https://medium.com/@annikabrundyn1/the-beginners-guide-to-recurrent-neural-networks-and-text-generation-44a70c34067f

[14] Brownlee, Jason. “Text Generation with LSTM Recurrent Neural Networks in Python with Keras.” 2022. https://machinelearningmastery.com/text-generation-lstm-recurrent-neural-networks-python-keras/

[15] Gupta, Purushottam. “Auto-generative texts from Shakespeare writing using Deep LSTM Recurrent Neural Networks.” 2021. https://medium.com/@purushottamgupta2010/auto-

generative-texts-from-shakespeare-writing-using-deep-lstm-recurrent-neural-networks-357253fe94a1

[16] Kant, Laxmi. "Poetry Generation Using Tensorflow, Keras, and LSTM." 2020. https://kgptalkie.medium.com/poetry-generation-using-tensorflow-keras-and-lstm-75c4e4b7f07e

[17] Sai, Chaithanya Vamshi. "Generating Poems Using Natural Language Processing (NLP) with Python." 2022. https://github.com/wozu-dichter/Poem_Generator_LSTM_NLP

[18] GeeksForGeeks. "LSTM Based Poetry Generation Using NLP in Python." 2024. https://www.geeksforgeeks.org/lstm-based-poetry-generation-using-nlp-in-python/

[19] Chaturvedi, Kabir. "AI Shakespeare NLP Driven Text Generation." 2024. https://github.com/kabirchaturvedi/AI-Shakespeare-NLP-Driven-Text-Generation

[20] Rajhen, Siva Vimel. "AI Poem Generator RNN." 2023. https://github.com/SivaVimel/AI-Poem-Generator-RNN

[21] Rivas, Daniel. "Shakespeare Text Generator." 2023. https://github.com/danieldotwav/Shakespeare-Poetry-Generator-Neural-Network

[22] GeeksForGeeks. "ML | Text Generation using Gated Recurrent Unit Networks." 2024. https://www.geeksforgeeks.org/ml-text-generation-using-gated-recurrent-unit-networks/

[23] M. De Araujo Possi, A. De Paiva Oliveira, A. Moreira and L. Mucida Costa, "A Neural Network-Based Language Model for Automatic Poem Generation," *2024 IEEE 20th International Conference on Intelligent Computer Communication and Processing (ICCP)*, Cluj-Napoca, Romania, 2024, pp. 1-8, doi: 10.1109/ICCP63557.2024.10792998.

Appendix A: Prior Work Architecture Comparison

Presented below is a summary of the architectural features of several poetry and text generation networks. Layer size is not included as it varies significantly between each project.

<table>
<tr><th>Type</th><th>Generation Unit*</th><th>Depth**</th><th>Other Features</th><th>Projects</th></tr>
<tr><td>RNN</td><td>Word</td><td>Deep</td><td></td><td>[3]</td></tr>
<tr><td rowspan="7">LSTM</td><td rowspan="4">Character</td><td rowspan="2">Shallow</td><td></td><td>[4] [9][10]</td></tr>
<tr><td>Bidirectional, Dropout</td><td>[11]</td></tr>
<tr><td rowspan="2">Deep</td><td></td><td>[12]</td></tr>
<tr><td>Dropout</td><td>[13][14]</td></tr>
<tr><td rowspan="3">Word</td><td rowspan="3">Deep</td><td></td><td>[3][15][16]</td></tr>
<tr><td>Dropout</td><td>[17]</td></tr>
<tr><td>Bidirectional, Dropout</td><td>[18]</td></tr>
<tr><td rowspan="2">GRU</td><td>Character</td><td>Shallow</td><td></td><td>[19][20][21][22]</td></tr>
<tr><td>Word</td><td>Deep</td><td></td><td>[3]</td></tr>
<tr><td rowspan="2">Transformer</td><td>Word</td><td>Deep</td><td></td><td>[3]</td></tr>
<tr><td>Phoneme</td><td>Deep</td><td>Based on GPT-2</td><td>[23]</td></tr>
</table>

* The linguistic unit into which the input text is reduced for analysis and generation.

** A network is considered “Deep” if it contains more than just the input, sequence analysis, and output layers (such as having a second analysis layer or a fully-connected non-linear layer).

Appendix B: Model Parameters & Training Data

All trained models shared the following architectural parameters:

Training sequence length	30 tokens
Data loader batch size	32
Embedding layer feature size	256
Recurrence layer(s) feature size	512
Fully-connected layer(s) feature size	128
Training epochs	20

Additionally, all character-based models sampled 100,000 input characters from the dataset, while all word-based models sampled 200,000 input characters.

Character-based RNN models

Character-based RNN: Baseline

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.33765	11.12319	34.565
2	2.05433	8.04246	69.289
3	1.94625	7.22715	104.135
4	1.87138	6.70338	138.884
5	1.81281	6.32599	173.622
6	1.76394	6.0169	208.73
7	1.72137	5.76442	243.837
8	1.68321	5.55992	278.945
9	1.64682	5.35725	314.09
10	1.61246	5.17298	349.222
11	1.58099	5.00485	384.283
12	1.5498	4.84692	419.282
13	1.51992	4.71534	454.299
14	1.49214	4.57778	489.138
15	1.46438	4.45237	524.126
16	1.43816	4.34518	559.024
17	1.41133	4.21959	594.012
18	1.38512	4.10989	628.976
19	1.35955	4.0075	663.948
20	1.33506	3.90847	698.885

Character-based RNN: Two recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.28061	10.47191	73.029
2	1.98587	7.52849	146.291
3	1.86604	6.66458	219.597
4	1.78218	6.14123	292.954
5	1.71313	5.72084	366.294
6	1.65265	5.38109	439.703
7	1.59733	5.10012	513.104
8	1.54519	4.83535	586.386
9	1.49621	4.59644	659.689
10	1.44794	4.38346	733.009
11	1.40139	4.1834	806.385
12	1.35524	3.98961	879.891
13	1.31068	3.81775	953.316
14	1.26442	3.64366	1026.694
15	1.21963	3.4839	1100.039
16	1.17499	3.32746	1173.46
17	1.12832	3.16932	1246.947
18	1.08349	3.02771	1320.289
19	1.03888	2.89477	1393.709
20	0.99274	2.76355	1467.093

Character-based RNN: Three recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	3	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.27094	10.35764	111.708
2	1.96692	7.38043	224.028
3	1.84054	6.51457	336.543
4	1.75005	5.9474	449.005
5	1.67325	5.50037	561.202
6	1.6048	5.13383	673.433
7	1.54231	4.83197	785.776
8	1.48106	4.53337	897.903
9	1.42414	4.28675	1010.192
10	1.36674	4.04935	1122.467
11	1.30974	3.80957	1234.869
12	1.2522	3.59993	1347.286
13	1.19635	3.39993	1459.839
14	1.13895	3.20942	1572.28
15	1.08099	3.0245	1684.636
16	1.02524	2.86076	1797.099
17	0.96635	2.68964	1909.769
18	0.90871	2.53843	2022.31
19	0.85228	2.39383	2134.889
20	0.79492	2.26025	2247.436

Character-based RNN: 20% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	20%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.37596	11.53157	35.545
2	2.10694	8.47796	71.241
3	2.00701	7.67771	106.862
4	1.94091	7.18549	142.515
5	1.88941	6.82441	178.273
6	1.84723	6.55823	213.932
7	1.81044	6.32322	249.668
8	1.77788	6.12149	285.319
9	1.74806	5.92706	320.928
10	1.72255	5.7881	356.609
11	1.69525	5.62553	392.273
12	1.6713	5.49713	427.849
13	1.64421	5.34096	463.484
14	1.62472	5.2392	499.127
15	1.60451	5.14013	534.789
16	1.58472	5.04623	570.366
17	1.5606	4.92307	606.051
18	1.54222	4.82945	641.651
19	1.52086	4.73288	677.2
20	1.49992	4.62613	712.873

Character-based RNN: 50% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	50%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.45937	12.54471	35.49
2	2.21532	9.46298	71.146
3	2.13046	8.6973	106.921
4	2.0734	8.22223	142.659
5	2.02974	7.87144	178.282
6	1.99022	7.57934	213.833
7	1.96261	7.36516	249.589
8	1.93297	7.15436	285.265
9	1.9098	6.98142	321.01
10	1.88521	6.81439	356.734
11	1.86649	6.69624	392.416
12	1.84707	6.56828	428.119
13	1.83043	6.45742	463.754
14	1.81007	6.32372	499.484
15	1.79457	6.21527	535.141
16	1.77925	6.1282	570.747
17	1.76187	6.03731	606.443
18	1.7462	5.93369	642.118
19	1.73091	5.84336	677.736
20	1.71537	5.75867	713.529

Character-based RNN: Bidirectional

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	True	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.33162	10.9633	69.409
2	2.07174	8.20192	139.213
3	1.96378	7.35906	209.044
4	1.88698	6.81923	279.085
5	1.82547	6.40264	348.9
6	1.77462	6.09135	418.666
7	1.72934	5.8165	488.47
8	1.68788	5.58744	558.209
9	1.64999	5.37177	628.302
10	1.61406	5.18124	698.551
11	1.58032	5.01214	768.353
12	1.54811	4.84785	838.026
13	1.51798	4.70556	907.691
14	1.48764	4.561	977.573
15	1.45941	4.43472	1047.424
16	1.43062	4.31437	1117.147
17	1.4036	4.19104	1186.906
18	1.37619	4.07302	1256.663
19	1.34806	3.96274	1326.392
20	1.32213	3.86151	1396.151

Character-based RNN: 1 Nonlinear Layer

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.36255	11.48553	35.189
2	2.05942	8.07541	70.364
3	1.95337	7.26967	105.036
4	1.88318	6.77357	139.67
5	1.82849	6.41282	174.487
6	1.78283	6.13244	209.203
7	1.74256	5.89794	243.907
8	1.70553	5.67343	278.582
9	1.6732	5.49134	313.266
10	1.64314	5.32776	347.957
11	1.61313	5.16648	382.638
12	1.5861	5.03029	417.45
13	1.56007	4.90385	452.138
14	1.53421	4.77741	486.782
15	1.50999	4.66541	521.588
16	1.48522	4.54493	556.289
17	1.46246	4.44045	590.918
18	1.44033	4.3436	625.618
19	1.41753	4.2444	660.466
20	1.39482	4.14572	695.242

Character-based RNN: 2 Nonlinear Layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	2

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.39314	12.04786	35.631
2	2.05756	8.07324	71.449
3	1.94663	7.2353	107.095
4	1.87348	6.72189	142.778
5	1.81504	6.34514	178.761
6	1.76578	6.03956	214.445
7	1.72255	5.78441	250.129
8	1.68234	5.55212	285.896
9	1.64674	5.35712	321.764
10	1.61163	5.18503	357.437
11	1.58048	5.00602	393.099
12	1.54908	4.85528	428.903
13	1.52051	4.72777	464.578
14	1.49049	4.58662	500.239
15	1.46424	4.45167	535.965
16	1.43741	4.33951	571.792
17	1.4108	4.22778	607.542
18	1.38427	4.12013	643.257
19	1.35983	4.02015	679.049
20	1.33401	3.91848	714.786

Character-based RNN: All Parameters

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	20%	True	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.30833	10.72407	175.341
2	2.04515	7.97694	350.827
3	1.94732	7.23224	528.439
4	1.87877	6.7525	705.332
5	1.82531	6.4075	882.938
6	1.78175	6.13888	1060.293
7	1.7416	5.89993	1237.615
8	1.70923	5.71483	1414.917
9	1.674	5.51012	1592.326
10	1.64582	5.3581	1769.654
11	1.61731	5.20391	1947.038
12	1.58994	5.05901	2124.221
13	1.56263	4.92985	2301.415
14	1.53905	4.80959	2478.528
15	1.51255	4.6903	2655.702
16	1.4873	4.56631	2835.247
17	1.46339	4.46184	3018.484
18	1.44194	4.36869	3199.541
19	1.41958	4.26366	3380.023
20	1.39807	4.17166	3560.508

Word-based RNN models

Word-based RNN: Baseline

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.90303	1412.0675	32.512
2	6.09897	506.12552	65.146
3	5.78791	364.06656	97.068
4	5.52008	275.86615	128.838
5	5.26725	213.40991	160.579
6	5.02432	165.53761	192.309
7	4.78392	128.60912	224.065
8	4.54861	100.85806	255.855
9	4.31439	79.81445	287.594
10	4.08276	63.08128	319.326
11	3.85338	49.99904	351.054
12	3.6269	39.64831	382.691
13	3.40537	31.67673	414.238
14	3.19097	25.63364	445.809
15	2.98285	20.7853	477.382
16	2.78357	17.06084	508.922
17	2.59169	13.97459	540.442
18	2.40629	11.60143	572.027
19	2.22864	9.67807	603.547
20	2.05727	8.16035	635.173

Word-based RNN: Two recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.77441	1150.24548	46.598
2	6.11522	519.06958	93.216
3	5.82451	382.56818	139.832
4	5.57114	292.91418	186.533
5	5.32214	229.4278	233.166
6	5.07704	176.30124	279.861
7	4.82598	135.98282	326.564
8	4.57336	104.54724	373.2
9	4.31638	81.07337	419.888
10	4.05514	61.61898	466.517
11	3.78971	47.18847	513.138
12	3.52423	36.10068	559.949
13	3.2584	27.55368	606.63
14	2.99757	21.10059	653.369
15	2.74107	16.32179	699.987
16	2.49092	12.67316	746.712
17	2.25233	9.93635	793.404
18	2.02239	7.87968	840.052
19	1.80298	6.29353	886.797
20	1.59547	5.08753	933.412

Word-based RNN: Three recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	3	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.75771	1103.67456	61.287
2	6.13568	528.84943	123.077
3	5.87519	400.42123	184.953
4	5.63884	315.49286	246.675
5	5.4145	248.8409	308.452
6	5.1907	198.01062	370.044
7	4.96696	157.25407	431.822
8	4.74082	124.19492	493.48
9	4.50629	98.45083	555.162
10	4.26928	77.29863	616.739
11	4.02887	60.13401	678.545
12	3.78315	47.14945	740.264
13	3.53445	36.50747	802.075
14	3.2868	28.44494	863.752
15	3.03794	22.12054	925.503
16	2.78963	17.19073	987.293
17	2.54832	13.41095	1049.17
18	2.31236	10.59095	1110.951
19	2.08395	8.40659	1172.694
20	1.86611	6.72237	1234.408

Word-based RNN: 20% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	20%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.87874	1379.07141	31.711
2	6.16855	546.07452	63.399
3	5.89356	405.84219	95.046
4	5.65095	319.68173	126.746
5	5.42411	251.20682	158.462
6	5.20633	200.10707	190.121
7	4.998	160.95239	221.813
8	4.78907	129.67203	253.449
9	4.57226	104.02588	285.182
10	4.37291	84.87315	316.829
11	4.17065	68.99534	348.429
12	3.96953	56.38702	380.089
13	3.7649	45.8712	411.791
14	3.57491	37.80024	443.455
15	3.38182	31.05318	475.108
16	3.19401	25.70527	506.761
17	3.01318	21.40605	538.354
18	2.8433	18.05751	569.978
19	2.67312	15.21111	601.574
20	2.50664	12.86938	633.291

Word-based RNN: 50% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	50%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.92555	1449.47107	31.668
2	6.32289	639.42719	63.334
3	6.11594	515.44489	95.03
4	5.93361	430.61429	126.754
5	5.76332	356.86804	158.43
6	5.60131	302.52475	190.033
7	5.43318	254.60858	221.632
8	5.27849	217.37708	253.358
9	5.12349	183.45547	284.977
10	4.97532	157.95233	316.641
11	4.82851	136.04581	348.33
12	4.67782	116.90759	380.023
13	4.52826	100.2788	411.64
14	4.38763	86.95816	443.193
15	4.23978	74.53548	474.846
16	4.09155	64.18365	506.514
17	3.95552	55.71796	538.153
18	3.81212	48.23978	569.843
19	3.6721	41.99846	601.601
20	3.53328	36.19315	633.25

Word-based RNN: Bidirectional

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	True	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.84336	1283.71277	58.192
2	5.95785	437.99487	116.715
3	5.55946	288.74994	175.087
4	5.20374	198.12579	233.421
5	4.8749	141.6405	291.644
6	4.56057	103.0317	350.011
7	4.25661	75.35023	408.454
8	3.96981	56.35825	466.985
9	3.69814	42.84553	525.294
10	3.44474	33.11816	583.815
11	3.21042	26.17469	642.131
12	2.991	21.01382	700.534
13	2.78437	17.0583	758.935
14	2.5868	13.96567	817.308
15	2.3981	11.51843	875.615
16	2.21748	9.61323	934.097
17	2.04264	8.03473	992.453
18	1.87759	6.80141	1051.02
19	1.71967	5.78696	1109.326
20	1.56782	4.94942	1167.903

Word-based RNN: 1 Nonlinear Layer

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.79125	1224.50818	31.74
2	5.93142	422.39044	63.501
3	5.57276	290.06668	95.288
4	5.287	215.64874	127.007
5	5.01894	162.27032	158.734
6	4.7423	122.62131	190.43
7	4.44794	89.96403	222.203
8	4.1359	65.94621	254.042
9	3.81909	47.7182	285.806
10	3.5151	35.20224	317.482
11	3.2446	26.96157	349.301
12	3.01189	21.34854	381.001
13	2.81168	17.47673	412.717
14	2.63423	14.59394	444.511
15	2.47433	12.42349	476.331
16	2.32624	10.71395	508.011
17	2.19082	9.31712	539.752
18	2.06267	8.2068	571.452
19	1.94203	7.26234	603.276
20	1.82762	6.46722	634.944

Word-based RNN: 2 Nonlinear Layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	2

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.94814	1434.72168	22.631
2	6.29492	618.58624	45.384
3	6.09625	508.3768	68.229
4	5.93317	427.06564	90.987
5	5.7861	365.55658	113.752
6	5.64931	320.24774	136.479
7	5.51753	278.26062	159.262
8	5.38797	244.13531	182.015
9	5.25545	212.92061	204.915
10	5.12222	184.74707	227.688
11	4.98471	161.30229	250.403
12	4.8459	139.24336	273.204
13	4.7017	120.73654	295.941
14	4.55202	103.74574	318.925
15	4.39474	88.04984	341.687
16	4.22913	74.51505	364.463
17	4.05789	62.78075	387.206
18	3.87626	51.84635	410.02
19	3.69152	43.06572	432.963
20	3.50525	35.48359	455.781

Word-based RNN: All Parameters

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	20%	True	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.68237	1072.61316	100.324
2	6.03821	474.6813	200.599
3	5.74728	349.48425	301.15
4	5.50367	270.99945	401.538
5	5.27702	214.44583	502.003
6	5.05716	170.44585	602.732
7	4.83597	135.92664	703.656
8	4.61392	107.83382	804.558
9	4.37718	84.65071	905.097
10	4.13256	65.75237	1005.671
11	3.87696	50.89915	1106.459
12	3.61407	38.94971	1207.069
13	3.35526	30.121	1307.895
14	3.11008	23.5557	1408.494
15	2.88789	18.82149	1509.124
16	2.68547	15.35186	1609.723
17	2.51534	12.92072	1710.225
18	2.36343	11.06736	1810.826
19	2.21487	9.54249	1911.205
20	2.08746	8.40049	2011.809

Character-based LSTM models

Character-based LSTM: Baseline

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.23234	10.12543	90.711
2	1.90391	6.91681	182.28
3	1.77807	6.10788	273.832
4	1.69138	5.59128	365.205
5	1.6231	5.23369	456.376
6	1.56537	4.93217	548.071
7	1.5158	4.69465	639.671
8	1.47036	4.48322	731.04
9	1.42979	4.30081	822.153
10	1.38988	4.13336	913.67
11	1.35322	3.98427	1005.03
12	1.31633	3.83238	1096.565
13	1.2806	3.70213	1187.911
14	1.24495	3.56481	1279.283
15	1.20843	3.43544	1370.654
16	1.17324	3.31406	1462.248
17	1.13672	3.19312	1553.772
18	1.10054	3.07781	1644.691
19	1.06334	2.96381	1736.241
20	1.02572	2.85451	1827.808

Character-based LSTM: Two recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.26374	10.48842	199.575
2	1.87563	6.73889	399.102
3	1.72943	5.82562	598.509
4	1.62882	5.25729	798.406
5	1.55276	4.86919	998.597
6	1.48714	4.55824	1199.009
7	1.42841	4.30534	1399.586
8	1.37447	4.07186	1604.696
9	1.32087	3.8543	1813.21
10	1.26788	3.65672	2021.848
11	1.2148	3.46385	2230.54
12	1.16002	3.2726	2439.563
13	1.10392	3.09232	2648.652
14	1.0452	2.91465	2858.352
15	0.98547	2.74336	3068.15
16	0.92118	2.57203	3277.917
17	0.85476	2.40144	3487.641
18	0.789	2.24373	3697.197
19	0.72004	2.09405	3905.427
20	0.65029	1.94802	4111.767

Character-based LSTM: Three recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	3	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.42958	12.54159	299.102
2	1.98577	7.54012	597.8
3	1.81723	6.36141	896.794
4	1.70036	5.65294	1195.521
5	1.60828	5.15193	1494.843
6	1.53147	4.77248	1794.22
7	1.46193	4.45484	2093.716
8	1.39738	4.17487	2393.278
9	1.3376	3.93276	2692.63
10	1.2795	3.704	2992.182
11	1.22256	3.49121	3291.471
12	1.16256	3.29035	3590.753
13	1.10373	3.09804	3890.819
14	1.04482	2.91931	4190.192
15	0.98327	2.7433	4489.711
16	0.92238	2.57736	4789.432
17	0.85824	2.41698	5090.538
18	0.79398	2.26154	5394.526
19	0.72737	2.11284	5698.782
20	0.66434	1.98119	6002.934

Character-based LSTM: 20% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	20%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.29731	10.86541	90.628
2	1.95983	7.31972	181.015
3	1.82283	6.38289	271.998
4	1.72566	5.7896	363.053
5	1.65157	5.38126	453.965
6	1.59181	5.06515	545.028
7	1.54056	4.81384	635.908
8	1.49358	4.58839	726.91
9	1.45171	4.40369	818.093
10	1.40856	4.20484	909.282
11	1.37102	4.04994	1000.158
12	1.33414	3.91123	1091.709
13	1.29643	3.75884	1182.993
14	1.2615	3.62673	1274.467
15	1.22763	3.51058	1365.279
16	1.19192	3.38605	1456.731
17	1.15915	3.27213	1548.212
18	1.1252	3.16525	1639.401
19	1.08726	3.04459	1730.524
20	1.05325	2.93852	1822.033

Character-based LSTM: 50% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	50%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.37004	11.74352	91.081
2	2.02689	7.82551	182.572
3	1.90059	6.90502	273.85
4	1.8111	6.30632	365.376
5	1.7442	5.91294	456.973
6	1.69192	5.60306	548.579
7	1.64516	5.35726	639.956
8	1.60337	5.12647	731.556
9	1.56522	4.93511	823.045
10	1.53083	4.76703	914.748
11	1.49737	4.60862	1006.356
12	1.46994	4.48698	1097.744
13	1.44064	4.35538	1189.393
14	1.41286	4.2351	1281.019
15	1.38528	4.11748	1372.744
16	1.35589	3.9986	1463.941
17	1.33181	3.90458	1555.767
18	1.30588	3.80326	1647.724
19	1.28113	3.70966	1739.431
20	1.25529	3.61448	1830.741

Character-based LSTM: Bidirectional

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	True	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.23877	10.20809	180.628
2	1.90909	6.95477	361.537
3	1.77128	6.06295	542.489
4	1.67205	5.48549	723.767
5	1.59395	5.07346	904.891
6	1.52712	4.74441	1085.886
7	1.46977	4.47609	1267.471
8	1.41707	4.24753	1448.373
9	1.36746	4.03601	1629.25
10	1.32126	3.85561	1810.454
11	1.27606	3.68502	1991.67
12	1.23247	3.52597	2172.571
13	1.18967	3.36966	2353.474
14	1.14619	3.22849	2534.775
15	1.10442	3.0912	2715.899
16	1.06112	2.96104	2896.789
17	1.01745	2.83168	3084.308
18	0.97505	2.71018	3271.991
19	0.93327	2.59684	3459.669
20	0.8905	2.48728	3646.75

Character-based LSTM: 1 Nonlinear Layer

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.36204	11.61518	93.605
2	1.99771	7.59084	187.283
3	1.86026	6.61301	281.21
4	1.76573	6.01589	374.937
5	1.69209	5.58948	468.629
6	1.63115	5.25202	562.057
7	1.5786	4.99112	655.229
8	1.53146	4.76203	748.742
9	1.4893	4.55954	842.337
10	1.45013	4.38233	935.852
11	1.4138	4.22587	1029.148
12	1.37915	4.07946	1122.845
13	1.34472	3.94148	1216.595
14	1.31254	3.81668	1310.053
15	1.28194	3.70172	1403.614
16	1.25078	3.58554	1497.698
17	1.22046	3.47954	1592.19
18	1.19045	3.3717	1686.478
19	1.16095	3.27285	1780.859
20	1.13093	3.17232	1875.12

Character-based LSTM: 2 Nonlinear Layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	2

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.43756	12.6746	95.697
2	2.04304	7.96645	191.57
3	1.89915	6.89438	287.574
4	1.80234	6.27154	383.617
5	1.72559	5.79021	479.317
6	1.66147	5.43246	575.474
7	1.60595	5.13743	671.525
8	1.55444	4.88388	767.791
9	1.50901	4.666	863.622
10	1.4688	4.48167	959.213
11	1.43061	4.30859	1055.155
12	1.39226	4.14934	1151.187
13	1.358	4.00649	1247.155
14	1.3226	3.86701	1343.245
15	1.28964	3.73862	1439.064
16	1.25621	3.618	1535.044
17	1.22426	3.50142	1631.043
18	1.19357	3.39119	1727.077
19	1.16149	3.28726	1823.035
20	1.12963	3.17665	1918.65

Character-based LSTM: All Parameters

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	20%	True	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.30756	10.8981	492.617
2	1.93766	7.15738	988.34
3	1.80542	6.27266	1484.305
4	1.71893	5.75782	1980.452
5	1.65404	5.38177	2476.464
6	1.59701	5.09002	2973.272
7	1.54789	4.83412	3469.787
8	1.50214	4.62778	3966.419
9	1.45965	4.43227	4465.349
10	1.41826	4.25408	4965.197
11	1.37853	4.0919	5465.611
12	1.33783	3.9225	5966.32
13	1.29834	3.76174	6467.02
14	1.25817	3.61395	6967.948
15	1.22027	3.48426	7468.96
16	1.17868	3.34037	7969.836
17	1.13819	3.2032	8471.587
18	1.0982	3.07262	8973.11
19	1.0561	2.94795	9474.699
20	1.0137	2.81897	9977.942

Word-based LSTM models

Word-based LSTM: Baseline

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.80902	1316.6864	56.842
2	6.07109	499.92255	114.104
3	5.77528	368.5141	171.141
4	5.50063	276.08331	228.525
5	5.23005	208.86736	285.414
6	4.96018	156.79623	342.8
7	4.68797	117.94933	399.794
8	4.41512	89.14381	457.042
9	4.13897	67.55432	513.897
10	3.86196	50.98428	571.198
11	3.58633	38.38396	628.23
12	3.31488	29.04743	685.409
13	3.04631	22.14775	742.493
14	2.78622	17.02149	799.748
15	2.53388	13.19213	856.841
16	2.29521	10.33092	914.032
17	2.06921	8.22421	971.041
18	1.85847	6.62867	1028.433
19	1.66188	5.43289	1085.402
20	1.47756	4.50209	1142.729

Word-based LSTM: Two recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.77933	1190.67651	102.294
2	6.14057	540.76917	205.181
3	5.94757	442.12729	307.529
4	5.77947	367.00397	410.289
5	5.61156	310.60434	512.944
6	5.43837	259.84845	615.587
7	5.25414	214.9722	718.237
8	5.06032	175.15849	820.368
9	4.85324	141.53139	923.474
10	4.63437	113.26092	1026.211
11	4.40357	89.19281	1129.226
12	4.16716	70.05516	1232.329
13	3.9197	54.24632	1335.589
14	3.66532	41.84176	1438.873
15	3.40796	32.19036	1541.688
16	3.14652	24.60723	1645.217
17	2.88852	18.99678	1748.577
18	2.63063	14.55191	1852.117
19	2.37895	11.3132	1955.43
20	2.13654	8.82939	2058.915

Word-based LSTM: Three recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	3	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.81548	1183.52905	148.303
2	6.18123	559.22223	296.629
3	5.99107	465.25528	445.346
4	5.83626	389.82047	594.462
5	5.68622	336.36969	743.485
6	5.53601	286.42377	892.491
7	5.38378	245.313	1041.641
8	5.22277	208.87	1190.858
9	5.057	174.85388	1340.045
10	4.88584	147.21254	1489.267
11	4.71112	122.48448	1638.622
12	4.53515	102.11801	1788.427
13	4.35859	85.12149	1937.904
14	4.17984	70.64621	2087.404
15	4.00416	59.51217	2236.831
16	3.83015	49.73463	2386.05
17	3.65284	41.42687	2535.161
18	3.47787	34.57599	2684.684
19	3.30697	29.04185	2833.894
20	3.13583	24.444	2983.219

Word-based LSTM: 20% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	20%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.79263	1333.38757	57.815
2	6.13425	537.09741	115.939
3	5.88624	412.80939	173.57
4	5.64654	323.11057	231.555
5	5.39656	248.36963	289.315
6	5.14533	191.41348	347.269
7	4.89425	147.13446	405.047
8	4.64459	113.92655	462.889
9	4.39674	88.17435	520.683
10	4.14353	67.86723	578.502
11	3.89648	52.84852	636.481
12	3.6537	41.18308	694.128
13	3.41574	32.24908	752.125
14	3.17828	25.39948	809.792
15	2.9518	20.19187	867.923
16	2.73095	16.09547	925.814
17	2.52029	13.02639	983.823
18	2.32021	10.62646	1041.613
19	2.12724	8.74008	1099.665
20	1.94578	7.25955	1157.435

Word-based LSTM: 50% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	50%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.80817	1388.92993	58.149
2	6.23068	588.40851	115.798
3	6.05974	492.66608	173.842
4	5.90431	427.41635	231.592
5	5.74582	361.77631	289.533
6	5.57665	304.0033	347.351
7	5.40245	253.30522	405.334
8	5.22889	211.10547	463.087
9	5.04365	174.19576	521.18
10	4.86251	143.32484	578.974
11	4.68207	119.09524	637.086
12	4.50509	98.89874	694.869
13	4.32688	82.89957	752.892
14	4.14719	68.20713	810.753
15	3.98224	57.68382	868.762
16	3.82012	49.11401	926.623
17	3.64729	40.92348	985.69
18	3.48435	34.73468	1043.561
19	3.33591	29.93388	1101.352
20	3.18357	25.67314	1158.725

Word-based LSTM: Bidirectional

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	True	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.74308	1237.93323	109.08
2	5.95152	444.51611	218.377
3	5.59306	303.004	327.636
4	5.24104	209.64844	437.048
5	4.89224	146.30011	546.26
6	4.54704	101.85964	655.304
7	4.20883	71.84274	764.687
8	3.87798	51.5364	874.063
9	3.5493	36.82804	984.078
10	3.23269	26.70235	1094.004
11	2.93256	19.80168	1203.798
12	2.64966	14.83563	1313.553
13	2.38578	11.39511	1423.42
14	2.14595	8.92795	1533.178
15	1.92211	7.10213	1643.048
16	1.71589	5.76243	1752.858
17	1.52542	4.74249	1862.446
18	1.34979	3.96309	1973.331
19	1.18877	3.36398	2084.422
20	1.03927	2.88605	2195.541

Word-based LSTM: 1 Nonlinear Layer

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.78382	1309.53589	58.419
2	6.08083	498.12357	117.008
3	5.81847	382.57495	175.871
4	5.59494	304.13535	234.545
5	5.37594	240.32649	293.311
6	5.14734	187.91483	351.657
7	4.90362	145.75218	410.217
8	4.64022	111.82141	468.225
9	4.35434	83.14834	526.542
10	4.03785	59.88177	584.576
11	3.69961	42.5214	642.757
12	3.35766	30.10874	700.719
13	3.04354	22.00728	758.963
14	2.77356	16.75485	816.775
15	2.54904	13.40055	874.993
16	2.35553	11.01746	932.731
17	2.18604	9.28193	990.952
18	2.0316	7.94462	1048.796
19	1.88919	6.87251	1106.928
20	1.75515	6.00418	1164.821

Word-based LSTM: 2 Nonlinear Layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	2

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.93984	1505.8999	48.213
2	6.24746	603.46594	96.987
3	6.07132	501.69592	145.417
4	5.94566	439.51654	194.302
5	5.81917	386.43201	242.857
6	5.69661	337.24759	291.431
7	5.58173	303.21667	340.262
8	5.46641	269.75095	388.762
9	5.34998	239.61084	437.654
10	5.23127	210.32576	486.235
11	5.11206	187.47865	534.841
12	4.99053	164.41464	583.78
13	4.86343	143.94749	632.313
14	4.7346	126.48916	681.221
15	4.59921	110.40547	729.785
16	4.45836	95.05843	778.586
17	4.31071	81.559	827.274
18	4.1584	69.74117	875.803
19	3.99681	59.18888	924.655
20	3.82807	50.13393	973.226

Word-based LSTM: All Parameters

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	20%	True	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.62557	1103.44971	249.904
2	6.10364	514.57977	499.982
3	5.91637	423.35132	750.197
4	5.75382	359.89197	1000.252
5	5.6063	307.33063	1249.88
6	5.46699	266.12912	1500.047
7	5.3311	232.17012	1749.955
8	5.19298	200.23538	2000.089
9	5.05377	171.63826	2250.29
10	4.9102	147.94841	2500.379
11	4.76278	127.445	2750.679
12	4.61482	109.56032	3000.991
13	4.45981	93.81983	3251.263
14	4.29992	79.37991	3501.701
15	4.12191	65.78941	3752.105
16	3.93406	54.10631	4002.157
17	3.7216	43.8331	4252.223
18	3.49534	34.85359	4502.445
19	3.26077	27.41568	4752.618
20	3.03349	21.74406	5002.606

Character-based GRU models

Character-based GRU: Baseline

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.26996	10.49298	104.454
2	1.92932	7.08983	203.212
3	1.80528	6.2656	302.64
4	1.72017	5.7516	402.427
5	1.65303	5.37794	501.695
6	1.59761	5.08201	601.632
7	1.54785	4.84038	701.189
8	1.50422	4.63014	801.724
9	1.46308	4.43789	898.354
10	1.42373	4.26612	998.047
11	1.38562	4.11012	1098.744
12	1.34953	3.96268	1199.346
13	1.3145	3.82184	1300.491
14	1.27778	3.68431	1401.243
15	1.24285	3.55571	1500.906
16	1.20722	3.43027	1601.705
17	1.17198	3.30953	1703.097
18	1.13574	3.1901	1803.321
19	1.0995	3.07369	1904.29
20	1.06204	2.95847	2005.438

Character-based GRU: Two recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.21701	9.99932	214.479
2	1.86212	6.6522	429.123
3	1.7301	5.82421	639.021
4	1.63772	5.29704	850.624
5	1.5634	4.92071	1061.431
6	1.50109	4.61475	1272.204
7	1.44424	4.36083	1486.367
8	1.39152	4.13507	1698.74
9	1.33829	3.91948	1911.706
10	1.28833	3.72732	2123.774
11	1.23617	3.53122	2336.895
12	1.18161	3.34022	2550.746
13	1.12736	3.16214	2761.074
14	1.07131	2.99148	2973.844
15	1.01283	2.81485	3184.369
16	0.95182	2.64479	3394.877
17	0.89074	2.4855	3608.831
18	0.82522	2.32668	3822.211
19	0.76176	2.17956	4035.437
20	0.69615	2.03791	4244.662

Character-based GRU: Three recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	3	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.2255	10.14876	323.063
2	1.85033	6.56577	635.054
3	1.70928	5.69631	944.883
4	1.61112	5.15382	1252.323
5	1.53472	4.77657	1559.802
6	1.46729	4.46739	1868.398
7	1.4063	4.19952	2176.368
8	1.34773	3.96008	2489.767
9	1.2901	3.73543	2802.367
10	1.23185	3.52089	3114.167
11	1.17075	3.31201	3425.002
12	1.10862	3.10457	3736.017
13	1.04402	2.90662	4048.228
14	0.97576	2.71021	4360.797
15	0.90401	2.52132	4673.243
16	0.83021	2.3414	4985.046
17	0.75268	2.15922	5299.675
18	0.67678	1.99749	5613
19	0.59875	1.84539	5927.531
20	0.52377	1.70923	6242.86

Character-based GRU: 20% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	20%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.30464	10.85607	98.739
2	1.97745	7.44639	198.489
3	1.85789	6.61112	299.141
4	1.78088	6.1198	399.393
5	1.72079	5.76717	499.673
6	1.66836	5.4676	599.715
7	1.62522	5.23302	699.791
8	1.58658	5.03145	800.328
9	1.54887	4.85505	901.748
10	1.51396	4.68049	1002.66
11	1.4826	4.53628	1103.133
12	1.44942	4.38408	1202.851
13	1.4189	4.24966	1303.21
14	1.38597	4.12165	1402.479
15	1.35626	3.99426	1514.478
16	1.32786	3.8858	1614.14
17	1.29694	3.7635	1713.72
18	1.26802	3.65199	1812.346
19	1.23615	3.5343	1912.247
20	1.20515	3.43079	2011.507

Character-based GRU: 50% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	50%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.40383	11.95312	98.422
2	2.07856	8.24152	197.762
3	1.96309	7.35292	296.873
4	1.89075	6.83194	395.956
5	1.83538	6.47026	495.091
6	1.78933	6.17851	594.483
7	1.75127	5.9578	693.262
8	1.71693	5.75569	791.718
9	1.68409	5.55445	890.571
10	1.65525	5.40359	990.888
11	1.62752	5.25243	1090.503
12	1.60406	5.13214	1190.334
13	1.57884	5.00807	1290.39
14	1.55596	4.89711	1389.898
15	1.53444	4.78063	1490.386
16	1.51008	4.67006	1588.947
17	1.48799	4.56906	1687.467
18	1.46154	4.45442	1787.155
19	1.43973	4.34984	1886.833
20	1.41432	4.24029	1986.389

Character-based GRU: Bidirectional

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	True	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.26587	10.34804	192.191
2	1.9377	7.14976	385.053
3	1.81135	6.30135	578.492
4	1.72368	5.77549	771.018
5	1.6552	5.39189	963.224
6	1.59826	5.094	1154.493
7	1.54663	4.83369	1347.226
8	1.50012	4.61169	1539.628
9	1.45798	4.41804	1732.993
10	1.41707	4.2417	1926.031
11	1.37898	4.07995	2119.477
12	1.34034	3.92282	2312.401
13	1.30332	3.78442	2505.328
14	1.26538	3.64153	2698.457
15	1.22821	3.50726	2890.725
16	1.1913	3.37631	3082.703
17	1.1537	3.24837	3276.233
18	1.11601	3.1243	3468.99
19	1.07783	3.00515	3661.464
20	1.03984	2.88974	3854.512

Character-based GRU: 1 Nonlinear Layer

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.34756	11.33055	98.743
2	1.98983	7.53009	197.561
3	1.85987	6.60595	296.278
4	1.77609	6.07493	394.845
5	1.71161	5.69798	493.91
6	1.65938	5.4092	593.217
7	1.61529	5.17519	692.837
8	1.57543	4.97098	793.619
9	1.54007	4.79928	892.708
10	1.50802	4.64231	992.415
11	1.47764	4.50248	1092.343
12	1.44876	4.37464	1190.895
13	1.42167	4.25646	1290.142
14	1.39508	4.14692	1389.541
15	1.36927	4.03835	1488.659
16	1.3439	3.93546	1587.423
17	1.31786	3.83241	1686.358
18	1.29412	3.74251	1786.161
19	1.26954	3.64945	1887.217
20	1.2455	3.55917	1989.24

Character-based GRU: 2 Nonlinear Layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	2

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.37719	11.93181	100.531
2	1.98296	7.49045	202.166
3	1.8549	6.58658	303.954
4	1.76961	6.05901	406.162
5	1.70601	5.67611	508.08
6	1.65106	5.37698	609.31
7	1.60587	5.13485	710.641
8	1.56457	4.92886	811.425
9	1.52701	4.74012	912.732
10	1.49207	4.58464	1013.66
11	1.45885	4.42659	1114.421
12	1.42839	4.30117	1215.779
13	1.39756	4.16001	1315.705
14	1.36885	4.04425	1414.691
15	1.33912	3.92449	1514.873
16	1.31184	3.8195	1614.772
17	1.28384	3.71259	1714.734
18	1.25569	3.60459	1815.842
19	1.22718	3.50257	1916.792
20	1.19862	3.4053	2017.001

Character-based GRU: All Parameters

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	20%	True	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	2.26625	10.40857	459.676
2	1.92862	7.09217	914.469
3	1.80723	6.27875	1367.928
4	1.72312	5.77674	1822.676
5	1.65717	5.40529	2276.295
6	1.60212	5.11243	2736.699
7	1.55601	4.89384	3192.946
8	1.5145	4.68336	3646.549
9	1.47837	4.51144	4100.432
10	1.44201	4.36024	4563.952
11	1.40671	4.20758	5026.231
12	1.37199	4.05227	5479.349
13	1.33899	3.92541	5940.275
14	1.30627	3.80378	6405.533
15	1.27669	3.68411	6873.247
16	1.24187	3.5556	7340.07
17	1.20894	3.43847	7804.755
18	1.18037	3.3433	8266.989
19	1.14737	3.22999	8736.013
20	1.11535	3.12822	9191.992

Word-based GRU models

Word-based GRU: Baseline

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	7.06373	1671.23706	62.465
2	6.26047	608.91992	125.231
3	5.93821	434.10434	187.751
4	5.64505	322.02509	250.352
5	5.35424	236.8989	313.704
6	5.06822	176.01463	376.559
7	4.78732	130.59097	441.807
8	4.50995	98.02908	506.54
9	4.23861	74.6459	571.424
10	3.9675	56.43793	636.438
11	3.70485	43.04913	700.965
12	3.44476	33.21796	766.312
13	3.19416	25.69913	831.515
14	2.95471	20.16547	897.612
15	2.7252	16.03925	963.066
16	2.50702	12.85559	1028.216
17	2.29801	10.39575	1093.563
18	2.09994	8.51262	1158.124
19	1.90898	7.00244	1222.372
20	1.72691	5.81238	1287.394

Word-based GRU: Two recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.98494	1469.61865	112.053
2	6.3002	641.39435	224.78
3	6.0313	487.42477	336.778
4	5.78363	375.49329	449.654
5	5.52307	284.479	561.756
6	5.25504	215.76398	676.674
7	4.98362	162.61157	793.831
8	4.71144	123.29109	908.921
9	4.43699	91.54778	1024.345
10	4.16317	69.63592	1136.32
11	3.8872	52.18093	1249.899
12	3.61045	39.50451	1365.073
13	3.33346	29.70857	1480.185
14	3.0633	22.58504	1596.157
15	2.79606	17.3043	1711.324
16	2.53714	13.30876	1826.432
17	2.28898	10.36484	1939.521
18	2.05123	8.10633	2054.768
19	1.82498	6.44216	2168.48
20	1.61031	5.17181	2281.798

Word-based GRU: Three recurrence layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	3	0	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.99463	1433.9884	152.618
2	6.32888	660.5658	303.105
3	6.08268	503.64862	453.514
4	5.85863	407.82709	603.786
5	5.6283	316.09598	753.178
6	5.39949	250.77257	904.342
7	5.16679	196.48938	1054.706
8	4.93107	153.75621	1204.876
9	4.6967	120.50019	1355.752
10	4.45925	94.05503	1505.346
11	4.22176	73.63708	1655.365
12	3.98238	57.99796	1804.52
13	3.74394	45.1571	1954.055
14	3.50708	35.75103	2104.448
15	3.27102	28.03764	2254.504
16	3.03865	22.10439	2405.257
17	2.80901	17.49697	2555.881
18	2.58867	14.01686	2706.066
19	2.37331	11.22888	2857.017
20	2.16529	9.07276	3007.032

Word-based GRU: 20% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	20%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	7.0411	1686.64294	61.868
2	6.3401	658.18378	124.193
3	6.07298	503.57733	186.009
4	5.81892	387.00351	248.174
5	5.56371	296.38074	310.088
6	5.31064	228.02304	372.37
7	5.05635	173.64159	434.393
8	4.81015	134.93817	496.566
9	4.56287	103.89558	558.307
10	4.32107	80.84913	620.772
11	4.07915	63.64739	682.584
12	3.85125	50.07138	744.942
13	3.62008	39.78676	806.569
14	3.39677	31.51178	869.313
15	3.18034	25.38402	931.524
16	2.97404	20.68691	993.89
17	2.77202	16.79306	1056.295
18	2.57912	13.84811	1118.18
19	2.40388	11.5773	1180.8
20	2.22917	9.68994	1242.921

Word-based GRU: 50% Dropout

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	50%	False	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	7.0484	1699.69385	62.437
2	6.47252	761.67279	124.566
3	6.28617	633.68085	187.002
4	6.1249	534.4231	248.782
5	5.95895	451.82022	311.49
6	5.79007	376.39685	373.309
7	5.61489	315.43234	435.812
8	5.44968	261.94022	497.633
9	5.27421	221.64413	559.376
10	5.1048	184.70517	621.031
11	4.93717	154.41846	683.053
12	4.76567	129.35431	744.862
13	4.59987	109.43724	806.995
14	4.4343	91.8355	869.328
15	4.2755	78.3424	932.174
16	4.11983	66.80191	994.377
17	3.96438	56.73412	1057.255
18	3.81293	48.44187	1119.788
19	3.66311	41.61803	1182.307
20	3.50593	35.64726	1244.301

Word-based GRU: Bidirectional

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	True	0

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	7.02354	1585.5575	119.733
2	6.15233	552.8197	240.343
3	5.7358	353.6705	361.944
4	5.33217	229.77969	482.245
5	4.9379	152.48282	603.202
6	4.55756	103.07835	723.895
7	4.19255	71.06213	844.465
8	3.84668	50.00608	964.991
9	3.52239	36.11761	1085.117
10	3.22755	26.87374	1205.223
11	2.96542	20.63301	1324.911
12	2.72838	16.14973	1444.921
13	2.50914	13.01335	1565.044
14	2.30126	10.49733	1685.129
15	2.10701	8.62412	1804.882
16	1.92028	7.12266	1925.058
17	1.74484	5.96075	2045.425
18	1.57458	4.99893	2166.054
19	1.41809	4.25801	2286.381
20	1.26543	3.65084	2406.308

Word-based GRU: 1 Nonlinear Layer

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.99119	1629.56409	62.622
2	6.09399	507.44418	126.354
3	5.74825	357.69263	189.249
4	5.46296	259.96301	252.69
5	5.18959	196.49959	315.524
6	4.90727	146.94604	379.094
7	4.60581	107.45904	441.597
8	4.27723	77.20884	504.522
9	3.92545	53.38674	566.845
10	3.56689	37.29786	629.701
11	3.23676	26.77929	692.335
12	2.95176	20.06997	755.541
13	2.70924	15.78655	817.872
14	2.49915	12.73296	880.695
15	2.31025	10.52878	943.317
16	2.14053	8.85623	1005.56
17	1.98557	7.59363	1068.313
18	1.8445	6.58379	1130.337
19	1.71285	5.74843	1193.033
20	1.58959	5.07245	1255.625

Word-based GRU: 2 Nonlinear Layers

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	1	0	False	2

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	7.14175	1815.26074	51.255
2	6.43974	751.37854	102.558
3	6.24805	612.03577	153.736
4	6.10011	526.64282	205.629
5	5.95541	447.52734	256.865
6	5.80718	388.22705	309.002
7	5.65307	329.15573	360.225
8	5.49755	281.5394	412.082
9	5.34203	237.44881	463.126
10	5.18783	202.79822	514.371
11	5.03041	172.3307	565.956
12	4.86995	145.11722	617.111
13	4.70432	122.38573	668.672
14	4.52806	102.57887	720.026
15	4.34296	83.69975	771.4
16	4.1438	68.64072	822.918
17	3.93167	55.12118	874.212
18	3.71041	44.27487	926.215
19	3.48483	35.19715	977.092
20	3.26247	27.92928	1028.743

Word-based GRU: All Parameters

Parameters

Parameter	Recurrence Layers	Dropout	Bidirectional	Nonlinear Layers
Value	2	20%	True	1

Training Data

Epoch	Average Loss	Average Perplexity	Elapsed Time (seconds)
1	6.83616	1313.77625	229.986
2	6.27068	616.3316	460.806
3	6.01823	470.18497	691.665
4	5.78683	373.43625	923.866
5	5.56052	292.96878	1155.471
6	5.33677	230.98729	1387.114
7	5.11713	184.11124	1620.255
8	4.89706	146.48112	1851.987
9	4.6781	116.57165	2083.178
10	4.45106	92.49575	2313.913
11	4.22093	72.63934	2546.398
12	3.97978	56.79623	2777.431
13	3.72253	43.58554	3010.085
14	3.45554	33.31989	3243.247
15	3.17952	25.10587	3475.125
16	2.9044	18.97272	3709.045
17	2.63986	14.57468	3941.671
18	2.39391	11.41194	4175.799
19	2.16358	9.03124	4409.527
20	1.96561	7.39547	4641.409

Appendix C: Output Text

Character-based RNN models

Character-based RNN: Baseline

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods and said If you could not see The passed and charms and ch
And on the pedestal these words appear	And on the pedestal these words appear'd The three montress and dares and said If you could not see The passed and charms and charm
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? There was an old person of Baracens and the sade and said If you could not see The passed and charms an

Character-based RNN: Two recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods Two old man of Sheen and the deep to the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the s
And on the pedestal these words appear	And on the pedestal these words appears Our heads are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? The water with a stream and horrid blue and best That departure from the painting by Cortona Whose deside the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat and so the streams are descend The deep the top of a hat a

Character-based RNN: Three recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods The sea-green hange And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sa
And on the pedestal these words appear	And on the pedestal these words appear'd Turn'd slack'ning the streams of the threl through the strange had to his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town And the sacred by all his sacred to the town
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? And what all the sea-shore armed him a traitor his head startled by a lovely sparkled fierce account of his shell And when the sea-green hands he cried In his hands are thing and sad To himself at his head so writing of my heart Who was stand The dark some she spirits along the sea-shore armed him a traitor his head startled by a lovely sparkled fierce account of his shell And when the sea-green hands he cried In his hands are thing and sad To himself at his head so writing of my heart Who was

Character-based RNN: 20% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood a look of Border of the bright and singled to the stoles and seat The head with his hands of the floor of his brows the towers and the propes of the dead so fair to his have the stream and the starch and she said A Quing of the grace of the prose is the bright be save the sand and said I follow and said I no more upon his heaven with the she to the store and said In old person of Bray and she made of the her wing to the have the son in see the song the was an old person of Sarken of the convers
And on the pedestal these words appear	And on the pedestal these words appearing and so she said If you may have stantly confessed the sight of the branks and the brint so stall soul from the face to his fate the sun rage on the town and sing he tore and grand and said An a thought so stood a more save and said I followed the sate and streaming the wave a felt the sight of the branks and said In the gods of the best and said If his shore in his hands of the back the lived to the round and said In old person of Borke was an old person of Buttle she sate and his head and s
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? O was a molthish and the streams and said If you be the said to the sat such a streams the store and shall soul and from the ward and said I will so the stream and sine of the stream to the streams and the streams are the right and shade of the three starmed in a stool and sad when the first the stars in a large rolling his head on a sing and see not the wroth and his hat he was an old person of Brands the head to sea The wave to see the wave the cave The best of the strange he touch'd the stre

Character-based RNN: 50% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods who cares of his foom the wint for menor of whose while the strange he went and the good the said An were he tound But they said The Cing on the side She propeets the thather of Stread and down and still he sear the sate the ware the slate and soon the sea she then the first of the shade She said and stone and liget so lady of Green who ears the stream'd and the stread and diston for the parnant see The plainted with a Crombon she were as the down and who went they water of my his bearth the s
And on the pedestal these words appear	And on the pedestal these words appearth-- On the carrous flow the charted the send the laid his hand her said I where the loved not the stromed the well and store and when they said I when the spired his dread of his by heart with his hands and his long when he said I was Paraceat was and sheart who was so they sook of his hath shee to the came the said I was and heart when the sumpenty day To his conseath the tower and strempled the sure in the singled they said he said I love the store and hears the streams and his has the fortin
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? O was a farther there and she the theer of the stread and come the shade And the sharn the said In the have his fore the painted he said I wall who land and land and shade her acked from the dead a Quing a hill And when the sould be formed with the see They lart of the shall speak They said I hear they said I who weat he said In the store he was an old person of II._ The said I wave I fan of his fouttons of the strong and steet the sea-- When the way of string her see There was an old person of

Character-based RNN: Bidirectional

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods of the dead of my long was strange him bread and said If you not the painting by t
And on the pedestal these words appear	And on the pedestal these words appear'd Head the people of the deep of standing by the painting by the pa
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? The bright be my long with the sea hath perplexed the brights the sea There was an old person of Strough the strange and brights the sea That expansing the people of the deep of standing by the painting by the p

Character-based RNN: 1 Nonlinear Layer

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood on a strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and
And on the pedestal these words appear	And on the pedestal these words appears the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the streams and strove to the
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? A was a little had som

Character-based RNN: 2 Nonlinear Layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods and said A Pobble who has not the silence the stars the st
And on the pedestal these words appear	And on the pedestal these words appear'd by the passed the stars the
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? The prince the stars the star

Character-based RNN: All Parameters

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods and the painting by the said And all the paint the town and so the sate to the first as he wear the painting the stares and the great was a hold be the Cid had to the chings said If who had shade the painted to the bear the servery made the said A many a fear And the mountain to the said To tear the sate the sate to the story and sungry beates and the shore of the story as the goes The close to the Charment of the sea the painting the town the painting the town And the great streams and the re
And on the pedestal these words appear	And on the pedestal these words appear the sate the three to the sea the feeling that the said If with the story and the terrease her sace to the town and the came to the sun and borne and the came to the streams and some shall said If will to the sky as the stream the mountain to the sate the sea The great was a border of the grasp of the shore of the day To the story and beggar and share the sat the strove the spark and the terrease he said If encerful fare the Cid to see the sea That the stranger there Whose heart and the count a
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? A said I be the seathen the sented the chirts and the charges therefore the town the Cid to the sea The words are dark and said A mate the fair thought the sea the sea to the sound a manner There was a young long the mons the sea to the Cranes the sea the wall and the shore of the Cid had not the Cid the said A many a great grows are the Charms and the thrones of the throus for the shade the three takes the stranger who carest and the canted by the sate to the field an old person of Bar. There

Word-based RNN models

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood To make our ways! and the Lust of Soul and The eager wind: I saw the God enter the fair-walled town And all the tilth of storm the nurse of him poured The other of Carthage the Gods begot her on the shore And bored out and both now with all that concourse was. And now amidmost of the hall and round about the shore And bored out their hands abroad and clave the udders' wave And now Ilioneus' stout ship her the bed and Gyas' sky And there he saw all sorrowful without the death-dues of the wine.
And on the pedestal these words appear	And on the pedestal these words appear and But when he saw the ships the Gods thereof in hand of war! And Tiber-mouth; fulfilled of all the leal of the Greeks And that chief charge of men and beasts of men and helms of holy But when he saw the twi-branched stem of sea and sea-cliffs of holy place The other lads with poplar-leaf and that he wrought? The Eastland of the Gods the earth-abiders' that she hath And now Aurora left alone Tithonus' saffron bed And all the tilth of Myrmidons: silent and all unrest And now Aurora left alone Tithonus' saffron bed
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? of her? The very shepherd Polypheme the sea men and men The Dardan folk to very a little space And all the tilth of Myrmidons: silent and raiment rolled A coal-blue cloud drew up to mine eyes all suddenly And there he saw all sorrowful without the death-dues of the wine. And joins the holy place the sea its lowest deeps upcast. The very whistling of the Gods the earth-abiders' that To bear the fire and the Gods and round about the shore And Phaeton's horses shining bright the ninth dawn from out the flint

Word-based RNN: Two recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood Such words he spake and from his mouth hand And now amidmost of the hall a banquet for her lord: And now Aurora from out his hands he saw the Danaans driving And he himself poles had that which the sky And with the naked a furrow exceeding by that gean the sea And there he saw the twi-branched that never ends When lo his father's image seemed to fall the weary about And all the houses of the sea and nursing with her eyes And with his father's image smit to meet the weapons' rain
And on the pedestal these words appear	And on the pedestal these words appear with The very shepherd Polypheme his feet doth skim And all the tilth is hushed and beasts and birds of many a hue; In thicket rough amid the hush of night-tide lay asleep And slipping off the load of care forgat its toilsome part. But as he reached his word he layed and the air; And they the Sibyl's dwelling-place and the dusk he went And all the houses of the day and feast he gives And all the unmeasured backs of them coil upon the sea Unto the sandy of the Gods against me with
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? that To Phoebus Ceres to tell all toil to meet But when thy deeds the sea a rock I never lay And now the looked-for in the sea its lowest deeps upcast. And now the looked-for day was come with simple light and sweet And Phaeton's horses shining bright the ninth dawn in did bear. Fame and the name who had the neighbouring people on the Now he who led his heart by wildfire of very door And I departing fain the hush and fair that she hath But if no less than my father then of me to meet:

Word-based RNN: Three recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood And ever sweet and Delos leaves and kingly quaked doth on. Then the sea by high of days or pity pebble And all the tilth of Priam's ship threesome of his God-maddened And many a thousand of cornel shafts amidmost of the foe. And called the Gods all overpassed but all the town of strife And first before their eyes of old doth cloudy And many a thousand there is facing host of dreams. For guilt of men and home of old and shook they war? A marble storm he of my beloved my life would never trow.
And on the pedestal these words appear	And on the pedestal these words appear of mine Cloanthus land And there the Queen of Drepanum and man of Rome. The very folk of father-folk the Sibyl's of men hath found Unto Laocoon; and to us what word he spake the land Of whom and roll he gives up and in his staring there: And now Aurora left upon the dark And now it all the walls to fare the sails we see And looking the bowls and the rest and brought the Sibyl's seat. And all the tilth of her folk and Greekish folk of gold And first neas of the Gods the holy house to strive
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? he saw the longed-for main and all the dimples world of gold; Shall from out the hollow of the Gods and men is bear their sails we lay And first of all the filthy of her arms and ancient heart And then her chariot: and so men by every side doth dwell Then close they fall for with such words and nothing rolled And many a thousand there is hung to the coming And portioning it is no fear to be that yet to me The Gods who be the ghosts of her arms and great a little band And called to come; of a slave: that I am be to see And bid her hand from heaven and in a happy to pass

Word-based RNN: 20% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood they fare; And first of times. is that they had wherein his hair And now Aurora from the house of Gods shall henceforth giftless go; And there neas gives to Italy and all the tides And all the host of the Gods the earth-abiders' of that foreseeing And now Aurora left alone Tithonus' saffron the sea And there neas with his hands and home to wake a mighty weight: And all the tilth of the Gods and men themselves of them The very whistling of the Gods and all the town beside. And all untended goatish flocks amid the herbage bite.
And on the pedestal these words appear	And on the pedestal these words appear And of the Gods and men and mak'st up are blowing. The Amazons of the theatre a mighty bull of strand. The mighty Juno's of Danaan ships and all the town in haste! And there the very lords of men and all the tides was tied: A black ewe thou hast found: and hollow wooded he cast. And all untended goatish flocks amid the herbage bite. And many a thousand men and bodies upon the sea the watery And now when all the Gods and men and all unrest And there he saw the twi-branched stem twin from the ground
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? of But thou the elder of the Gods the earth-abiders' of the earth. And with the wind the heaped-up waters drew the flame of yore The very shepherd of Troy and land that she was five And on the Cyclops' strand the sun on high set down. And with the wind of the sea and sky. milk to bear And all the Gods of her folk and folks the very town: And therewithal as they went of old and Death And then the ships of haven shall I nothing know: my ways The other ones with all my heart in a rock-hewn cave

Word-based RNN: 50% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood And witless hear'st and length the walls of godhead And now amidmost of the Gods and round about the heap. And now the Father neas on the shore against the sea And first the very bones of gold and gems the deep And many a mighty shade at last the noisome of old And now the very bones of the men and bodies of the strong? And now he left his heart and all the town to him And now when lo the holy doors of old shall fare. And thence away
And on the pedestal these words appear	And on the pedestal these words appear And ever the haven of the sea and all the holy doors And therewithal as tells the tale and feast alike therewith And with the wind the stars in the sea leaps the sea Now wendeth of the Gods beside on oars the shore And so Buthrotus' city-walls and the sea to tell the fire And all the unmeasured of the sea and mid the dusk flame And now the haven of his folk and voice of all And when the haven of the Gods and round the altars clung: And now the Gods of men the Gods of men and die
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? But Venus an hundred fatted to the Sibyl's of the sea And there ran he himself was the first of all the concourse And now the Father of his folk and on the wind And there the holy water there and so great; son And all the haven and the very corpse fair the breezes And then the haven of the sea and mad with mighty hand And therewithal a mighty oak is left a day to come But when the Father of men and beasts the holy place of old And now the haven shipless and call the altars of the foe.

Word-based RNN: Bidirectional

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood So sayeth a portent terrible with thee everywhere; O wretch The ancient mother of the Gods bound round the holy Maid With sail and oar the Queen of God hath no eyes The other spake: Apollo's shrine the secret stream anigh He spake: him the honey-juice with poppies bearing sleep Whereby she maketh shift on tree the hallowed bough to keep. And now amidmost of the world of God and all the sea Than that which the very feet and orbd shield of dread; Then through our quaking hearts indeed afresh the terror spread
And on the pedestal these words appear	And on the pedestal these words appear and clave the matter long. And all the unmeasured backs of them coil upon coil The very haters of the sea Orion's watery ways So the gleaming weapons there She dropped from her breast and all the gift-giving Quoth he: If any valour hath or heart that may withstand But thou not see with guard to watch that I grow But if somewhere is gone away from Troy the city of the earth. And as the South bids the sea and all the tides doth fly Then on his shoulders clang the shafts. of her heart:
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? in Now while the rivers seaward run from out the South: And as the South of the Greeks and nursing fields of doubt. Then straight neas caught away from the way abode. And all amazed stands moving nought with him with her golden imagery; The very whistling of all men and all the sea uptorn But now Grynean he himself his word he layed the Cretan shore.' So saying she blew the earth and sea and make a fish-cake. And now amidmost of the world of God and all the sea Than that which the house of oars and bows' three-headed push.

Word-based RNN: 1 Nonlinear Layer

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood But after they had fared with the twain and cups And merry of the Gods and round the holy doors undone; And there he saw neas' ships o'er the night-tide sea And now Aurora falls the tomb methinks of the land of the And there the tangled house and innermost and all the world of Rome! But when the cot-built of the Gods and ask and all the tides of storm Make thou not see the Myrmidons' that she sendeth the sea And witless hear'st the Gods and all the gift-giving of the Greeks; Whether with fraud and all the world of the Gods invite And now Ilioneus' stout ship and all the gift-giving of the Greeks; Whether with fraud and all the world of the Gods invite And now Ilioneus' stout ship and all the gift-giving of the Greeks;
And on the pedestal these words appear	And on the pedestal these words appear But now they say the sea and slain the sea and high And now Aurora from the heavens and the fight of the ways. Some choose the fire the sea and slain the sea and And now Aurora from the heavens and the fight of the ways. Some choose the fire the sea and slain the sea and And now Aurora from the heavens and the fight of the ways. Some choose the fire the sea and slain the sea and And now Aurora from the heavens and the fight of the ways. Some choose the fire the sea and slain the sea and
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? And why they sought the Gods and the Gods to aid And all the miserable of the Greeks the sea and slain And now Aurora from the heavens and roar the sea And witless hear'st the Gods and all the gift-giving of the Greeks; Whether with fraud and all the world of the Gods invite And now Ilioneus' stout ship and all the gift-giving of the Greeks; Whether with fraud and all the world of the Gods invite And now Ilioneus' stout ship and all the gift-giving of the Greeks; Whether with fraud and all the world of the Gods invite

Word-based RNN: 2 Nonlinear Layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood And all the Gods of Italy and all the Gods was I knew to be And then a mighty cloud of men and all the sea And therewithal as he saw the Gods that never ends And many a mighty bull of men and all the sea And therewithal they sought it and all the sea and fold And all the Gods of all the Gods that mighty hand to be And now Aurora as a little orphan and all the shore And then I pray thee first and all the sea uptorn And many a mighty bull of men and all the sea
And on the pedestal these words appear	And on the pedestal these words appear the fall. folk And all the Gods by chance a little space of all unrest And all the Gods of Italy and all the Gods was I knew to be And then a mighty cloud of men and all the sea And therewithal as he saw the Gods that never ends And many a mighty bull of men and all the sea And therewithal they sought it and all the sea and fold And all the Gods of all the Gods that mighty hand to be And now Aurora as a little orphan and all the shore And then I pray thee first and all the sea uptorn
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? with But when he saw the holy house and all the sea he said And many a mighty bull and kisses of old tie. And all the Gods of Italy and all the Gods was I knew to be And then a mighty cloud of men and all the sea And therewithal as he saw the Gods that never ends And many a mighty bull of men and all the sea And therewithal they sought it and all the sea and fold And all the Gods of all the Gods that mighty hand to be And now Aurora as a little orphan and all the shore

Word-based RNN: All Parameters

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood And now meseems the cruel rock we fare and all the town beside. And now Aurora from the lofty deck meanwhile assured of flight And all the heavens of the Gods and memory the watery way? And there three times about with hollow shade and hidden gainsayeth And now the looked-for day the Gods shall stay to be kind And now Aurora from the heavens and feast the guard of holy And now the Gods to aid and all the heavens fulfil. And all the Gods to aid and all the mountain steep And there neas gives Sergestus promised gift in hand
And on the pedestal these words appear	And on the pedestal these words appear And now the bowls of the Gods and memory to utter anger And all the Gods and bare the Gods shall stay And all the miserable of the Gods and men of the war: And there he saw the sea leaps and all the waters cast: And now meseems on the lips and the shades of right hand And now meseems all the Gods thereof did he hale The very shepherd Polypheme and the dusk moon did shroud And then the inmost realms thereof with the dust the grave And there neas chief a bale and all the town beside.
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? And all the Gods and memory the stars aright on the lover And now he sees the word ere with the wet thereby And there three times about her neck and stays the cares And there the hill of the Gods and memory to utter And now meseems the Danaan fleet of his ship beneath the night! And now amidmost of the Gods and ask of the horn And there he saw and ancient house and all unrest And there three times about with hollow shade in the air And now perchance and the hoary Faith Quirinus and all the town

Character-based LSTM models

Character-based LSTM: Baseline

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood as there And so the sea and there the holy strand And strive the winds and there the holy strand And strive the winds and there the holy strand And strive the winds and there the holy strand And strive the winds and there the holy strand And strive the winds and there the holy strand And strive the winds and there the holy strand And strive the winds and there the holy strand And strive the winds and there the holy strand And strive the winds and there the holy strand And strive the winds and t
And on the pedestal these words appear	And on the pedestal these words appears the sea the way to be The son of soul to be bore and leave the sea and therewithal the thing they seen And there the mighty love the deed so fire to see the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to le
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? What man for thee the winds and let the sea the way to be The son of soul to be bore and leave the sea and therewithal the thing they seen And there the mighty love the deed so fire to see the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the way to learned the sea the w

Character-based LSTM: Two recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood about we went out within the ships and sought they fill to bear And give her all the sea and wild and speed amidst the threshold of the land that stone upon the shore; And his will I sea your ears have forth a flood that wasting place The eager seamen and a watch wed sought Then through the handy southed in the war Sent by the flood and first of holy blood deserving ships and sought Then through the hirs son of the land that stood The hangings of the Gods and men and field and field and field a
And on the pedestal these words appear	And on the pedestal these words appear And great shoot the land that should she was all the sea and wind the handing stand Then lo the holy tripod shores at last that hand the shores of that Troy bear And give thee forged to the house there on the war And striving wide away from Troy and fortune fell of forsooth. And therewithal the handicraft and field and field and field and field and field and field and field and field and field and field and field and field and field and field and field and field and field and field and field an
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? With an arrow wondrus there And by the flood of folk and folk for course was turned And fear from out the house there is the word of Troy and fortune fell of from and for all did fleet with folk and sought the herp of heaven and on the war Sent by the flood and first of holy blood deserving ships and sought Then through the hirs son of the land that stood The hangings of the Gods and men and field and field and field and field and field and field and field and field and field and field and fiel

Character-based LSTM: Three recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood. Why should the door be shut? The faithful follower of one one of the crowd; Then rose and loudly to the coursers heart. Thus as he spoke in death have borne And in the son of man the son of God Weeping and wailing of men As my son while a little from the shouts; And burning with his hand hath somets heav'nly brighten The force of our son of man the dead. Then throws of honour mes; There was still and bade the sea of the sea-- Why should the door be should strong The storm-swift Iris will you c
And on the pedestal these words appear	And on the pedestal these words appear'd The horses of the sand and there Appear'd the watch beside him The together of a son of Goddess that hath should heard the desert calling and my heart and shouts The comrade of the sand and horses and the sea the sons of Greece And halted thee the coursers as herms with herden they laid the dead And pond'rous mass as they treasur'd; and on the sand And par'd to the shore of Goddess though dragg'd along the steeds And pass'd his comrade of the sand And pass'd his recost to thee they steps atte
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Had deen by the heav'nly Guide o'erspread The treasures and the sea of Goddess that hath should heard the desert calling and my heart and shouts The comrade of the sand and horses and the sea the sons of Greece And halted thee the coursers as herms with herden they laid the dead And pond'rous mass as they treasur'd; and on the sand And par'd to the shore of Goddess though dragg'd along the steeds And pass'd his comrade of the sand And pass'd his recost to thee they steps attendant one the steed

Character-based LSTM: 20% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods he spoke: The took the strength the spring The fair of the ships the rest That thou thy sons has spead; And then the window benow to rest The proding hand he sign'd and to the cars And but the many and the man and smortal came The mares and then the world and mind was spoke: The rose--will but the many and the chiefs of man In answered the mornals close he took the pyre the mornals close With prizes strings to sent the mores And there's that they son of man his seep And the hand and but the mi
And on the pedestal these words appear	And on the pedestal these words appear'd The many and the man and far and fear The tent they stood and from the course His founted and the sea the sons of Greece The many and the sight hath been stood And forth whoe'er thou with the ships the corpse of string And the son of Achilles swift of foot The comrades brought the sign and the ground With the shall be on the sign with the stars-- The mind come back to the man and spedded strife The car of the son of Thating counsel plac'd The monarch's son his sons his prizes and the senses h
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? The first to thee the charioteer The missing on the plain and the shore Ald man in days they laid his side. The fire they stark-but share the morning car And the son of the sprang of the sun was spoke: The comple pain and to the sea the storm-shiply thee The mares and the charioteer the day The first to the shores of Troy To whom the sons of his single remast That forlower of all the garden sear'd; And the there of the shores of man The course and brought the stars; The first to the son of Troy

Character-based LSTM: 50% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood And shall the stream thy horses and the stream shall come to thee And the word and the was the Gods And brought the world and the reach'd and straight A done the door of the ransom of the car And to the camp and the cars the car of strongest strike the horses and the course The cordses of the corpse to rest And in the door of the ring the world and shall the stream And the both and strong the sea the race of Greece And on the ships the glore-- Well be the mules the stream the dark-prow'd string
And on the pedestal these words appear	And on the pedestal these words appear'd And loudly to the car and the heart and the return. The fair-foolow'd on the hand he reclos'd. Then three to the rest nor on the plain. The conty stars the world when the world And when the many of a golden strain The course and suppliant himself the ring The strain'd the dark the roses strength And fasten'd to the ships the car The rest when the rest returning far And bore the dost the bread and son the doon the door And with the sent the world and me as the condur stream the dead. The corps
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? The rose of the wind and with his sing And loudly there and son of Troy The corpse they rage of fire And body on the store-- My hear he stood the rose and the breath the Gods and spread with the hand his car And son to me the dust; but when the wain and the rest And with the world and when the wood was sea And the was the share the wool-pointed strang The corpse to godlike sire the city share the car And son the corpse and thus the Grecian ships The world when the son of man and son And many th

Character-based LSTM: Bidirectional

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood. Then thou art left me now before the dead. The prizes of the stands and the stream He said; and straight his sons had plac'd And sunder in the ring the store And pest'd the mare they came they laid And to the ships the storm-swift string And there the door of the store Of life and more the meal may steeds And o'er the sea the storm-swift string and there the dead. The prizes of the stands and the stream He said; and straight his sons had plac'd And sunder in the ring the store And pest'd the m
And on the pedestal these words appear	And on the pedestal these words appear'd The mourn and wealth and straight And leading out the court are laid The storm-swift of the stor-- Why should the door be strain The longers said; and with the store And the way of the stor-- Why should the door be strain The longers said; and with the store And the way of the stor-- Why should the door be strain The longers said; and with the store And the way of the stor-- Why should the door be strain
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Have longly to the stream to drive And when the stood of men and men The stood and loudly chief the door The storm-swift with the storm Of shouldst standing all the dead. The port and loudly to the chiefs of God And sought her son of men And there was many a some words And there was many a some words And there was many a some words And there was many a some words And there was many a some words And there was many a some words

Character-based LSTM: 1 Nonlinear Layer

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood of the stream thou thy sons of all And a golden came the wind of the stream thou thy sons of all And a golden came the wind of the stream thou thy sons of all And a golden came the wind of the stream thou thy sons of all And a golden came the wind of the stream thou thy sons of all And a golden came the wind of the stream thou thy sons of all And a golden came the wind of the stream thou thy sons of all And a golden came the wind of the stream thou thy sons of all And a golden came the wind of
And on the pedestal these words appear	And on the pedestal these words appear'd The comrade of the course and straight and there he stood And there a lonely of the course And the will be old man and the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we se
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? The said and love the world and the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And there we sent of the string And ther

Character-based LSTM: 2 Nonlinear Layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood and the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou
And on the pedestal these words appear	And on the pedestal these words appear'd The harriors of the ships they stood And there in his sons the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the strin
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? There was the speed and son the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the string And there's thou the string the

Character-based LSTM: All Parameters

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood And stream'd the spirit of the first And she who said I saw the sun At the first circlet to the story which they won The saintly spirit of the first With this endure that thou hast seen The world and lively to the sounds Of the great shine of the first And leath the sun who had thee The sun desire contention on the seath or sorrow she said The first that see the sun when the third Its grace that thou shalt see the sound Of this heaven shall seet them thou shalt forth The shape which thine eyes
And on the pedestal these words appear	And on the pedestal these words appear'd to the starry which thou seest The chill of the seep of this deeps The countenance when thou shalt thou seest that seetneth there Which thine eyes are the sun when the soul Who speaks the sun where the soul Who such a speeding love that doth not like to thee The first thine eyes and the word Is brightness of the sun delight That in the seven shall be the sun A little shared to the first that first the soul Of the fifth or mortal world she said I saw the soul Which counterse the depertion of t
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Where the spirit of the sun delight That thou my single as the sun With greater childing and the sun And saw the sun where the sun destreted light Was the semblance that the starry cross Within the sun where the sun declare In this end which thou mayst thee spake the sacred splendour or deed And the first came that thou that the see When thou shalt seen and thing Their circle they who had the truth Which thou mayst seeming the truth I way the several straight of the seether which the shoot Shal

Word-based LSTM models

Word-based LSTM: Baseline

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood Of the Calfucci still of our impictur'd And the great Harry he who in the van And in the world was dragg'd the world Of that great Harry he who in him Hath ever and he one who hears And of his own sufficiency The saintly soul That thou mayst see the other lights to see That which the loftier which thou hast taken in thyself To be surpass as in that clime is
And on the pedestal these words appear	And on the pedestal these words appear resuming of God op'd op'd I beheld For this ladder is none the world Of Joshua who in his mirror spies The virtue of his spring-head: and love With such effectual working and that good Which from within the pleasant cloister's I To me who in this heav'n is none It is that in its state of love But is that ever in its seed Trangress'd the virtue which it knits and made
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? ken Of the great Baron he who in him and him And he was not yet so long a sign Of him who sets his return The saintly folds and his glory: The column cloth'd with verrey still was seen Unshaken: the Sacchetti still were great With them who blush to hear the bushel nam'd. Of the Calfucci still the branchy trunk Was in its strength: and to the curule chairs

Word-based LSTM: Two recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood whom The saintly circles in that happy world Whose base affection bodies: whose course forth Until which the first Florentines. The truth That in the circle circle of the world That in the world of all our life The other which in their tourneying it Such as they were in itself and conforms As one who in a mirror of his skill. And in that upper was many such was
And on the pedestal these words appear	And on the pedestal these words appear my sun That I have answer thee that thou dost know To thee the first whence thou hast hast seen To thee that with its regal enlarg'd The other rapture and pommel in their house. Was had been fasten'd on the point that he The world's deceitfulness Then him who was the soul That on the world of her own light should rise; Is I have spoken thee it henceforth is To whom a pilgrim light thou shalt then
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? That he who in their own modern phrase blind With the eye's where flitting and Giuda And Infangato they had to its particular me So that they to me of thy own and That thou mayst see the truth Of that which thou dost not give thy mind For I have thee to thee who thou hast inquir'd Thou hast thee that which thou dost henceforth But for its own eyes what he was on

Word-based LSTM: Three recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood of mind Which thou mayst see thee clearly to thee To Beatrice as I saw thy mind And they who mingle in this thoughts It is to marvel their feathers The spirit of heav'n who sworn the sky. Of the great Baron that swift now with day Their plac'd but so new I saw To this his will which thou may see To Beatrice as thou may not thee thy words;
And on the pedestal these words appear	And on the pedestal these words appear made out To Beatrice it who be thee hath be To Beatrice as thou may not thee it up To Beatrice who thou art not thee to hear I saw not no more of. and truth To other assemblage! thou may be thee not see To this thy mind when thou hast see I choose To Beatrice who no wish Head To this thy mind may be not told thee To Beatrice as thou may not thee to hear
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? And from the light that vests me I To put such a glory of this Which not me within me in such part The other semblance; which Tell me To one triangle in other thoughts The other spirit of whom this height. From such who borders an vaunt eye And from that verge as they may not choose For Beatrice as I saw the sun

Word-based LSTM: 20% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood Expect stars; with his eye The path that which from the first love rolls The universe of God. O patience! I spake And of the fluent phrasemen make your king; From your sublunar world to fight e'en as clear The universal form: of his furry The sacred chords that in the bosom of heav'n Hath satiated and another one of his own: The sacred who in their own land The o'erweening brood That plays the dragon after him that flees
And on the pedestal these words appear	And on the pedestal these words appear and I beheld The dame of angels who sang the And as the glow burns ruddiest the proudest summits; And to the circle of our city The ken of the angels deck'd I saw The fountain of their dew. In the dance That of the angels which in the world and fram'd And every line And with the beam Of that which in this heav'n is none And in that region where the sun
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Tread And if that I had spoken to thee That hath my wish for my ken was I yclept: And thus with such figuring as I believe The fountain whence your arts derive their streame. With such as a space at his father's love. Who for his own discretion Then saw I turn'd If thou hast nam'd of my side to And whence they hither came more honourable It is to pass in silence than to tell.

Word-based LSTM: 50% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood Of the Calfucci of the M. I tell Was the great of that which the world Of this life and he who sang The spirit with his seed; The will I look'd; That hath in his audience already keen The other worthily from the Holy Spirit The city's and the rest of the truth That all the world of that which I saw me With the great Baron of the world
And on the pedestal these words appear	And on the pedestal these words appear dust To raise the warp of whom grace And as the sun doth the galaxy from the light And to the summit which thou hast me And visited to the world And in his own and love to hear The grace divine which thou dost me to hear And put a period to the mortal wax Who to the throes For that he mead Of his bounty and he had his flock
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? And as the saintly of that which thou dost The other trine that falls the world Of that treat whose of the world The walls whose keeping The Conti were possess'd That plays the strain; of the world which thou The body's frown And as the moon Of the great Maccabee is the world The festival and the rest of their own They yet the Gualterotti of our wishes

Word-based LSTM: Bidirectional

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood and size And if thou hadst it? But since I see To thee the first Florentines. that which thou dost Is in itself In order each As if thou shalt see the mark and see of God The greater in that streaky fire. The realm But for the peer Orlando which thou hast inquir'd And for the great shield wherein he is To save unto the human nature lays The rest were complete But as thou seest
And on the pedestal these words appear	And on the pedestal these words appear who feeds And entertain to this upper that our wishes The first of our vows and through there The yellow lilies and with partial So that it bears the mighty tide Of my great offices and believe be laid The saintly creatures flying sang and in So suddenly that which thou dost believe But since hath disaccustom'd and from the rest And as thou shalt see the air in heav'n
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? give And thus the saintly orison began. To whom the keys of this heav'n is The world possess'd and through there replied: Who draws the cause of his own ways lead The fondness that stirs from the day when That onward thence along the sun As from the virtue dies and the fire That hath made wander both the sheep and lambs The sheep is Then parting and that

Word-based LSTM: 1 Nonlinear Layer

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood the Of that idea which the first trine ends. I saw The that which from the world cannot to smite The path that which from the world imprints The Lady of that incense I was heard: The brightness of the Supreme which views The sacred stairs that which thou dost not Of the mind's impress; that which thou behold'st The other points and in that array The saintly creatures flying and so far
And on the pedestal these words appear	And on the pedestal these words appear to And as the arrow is the snow unseal'd; That save in the air expatiates it was And in the semblance of that rain The pair of his German shores. The o'erweening The dame of that which views from the world Whose base affection many a spirit soils And as the arrow of the cord light Of Joshua that stain'd which I saw The sacred stairs of that incense I cried
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? I saw And as the arrow of the world imprints And I pray thee that thou dost thy words And in the spleen of that unerring bow Which I pray thee to that which I dwelt The saintly creatures flying sang The other points and in that array The saintly of the world proceed And clearer that which I had not to my And leaves that which thou dost not to see

Word-based LSTM: 2 Nonlinear Layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood The other triumph of that which thou seest The other triumph of that which thou And as the first which thou hast in heav'n And in this heav'n is in that which thou seest The other triumph of that which thou And as the first which thou hast in heav'n And as the first which thou hast will And in this heav'n and what I saw I saw And Infangato O everlasting of that
And on the pedestal these words appear	And on the pedestal these words appear who That Ubertino to thee which thou hast not for thy toil. The other is that which thou hast in heav'n And as the first which thou hast in heav'n And as the first which thou hast in heav'n And as the first which thou hast in heav'n And as the first which thou hast in heav'n And as the first which thou hast in heav'n And as the first which thou hast in heav'n And as the first which thou hast in heav'n
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? and I The other nature that which thou hast write And from the first which thou hast wish And from the world I saw I saw I saw And Infangato O everlasting of that And I had not the world I saw I saw And I saw that which thou hast mortal joy And I saw I turn'd me that thou hast joy And from the first which thou hast wish And from the world I saw I saw I saw

Word-based LSTM: All Parameters

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood And as the saintly orison that draws the sun Of sore repentance and Signa's the wave Of that which I tell thee from day to step The festival of Thomas lore And as the summit of that saintly lamp The other cause: and in degree of bliss And thus it equaling of the sacred book Of that prime counsel wherein the world Of that region and the world proceed
And on the pedestal these words appear	And on the pedestal these words appear Of that celestial appearances from the world Of that region of the world The other cause: that which thou hast inquir'd The fountain of the world of God shall sit And I am turn'd me to the ground; Of that perfection of the Holy Spirit The house of the world that which thou seest The other points that soar Forthwith the sun Of Aguglione's hind and the human join'd.
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? The County Guido The Conti is near The other points that soar with plumes And know'st and Bostichi. At the cross That is the semblance that which maketh The whole. The city Spirit is made That in the sun that which thou dost report The County Guido and whoso in her And as they hither in the world Of that region The Conti is grown

Character-based GRU models

Character-based GRU: Baseline

<table>
<tr><th>Seed</th><th>Output</th></tr>
<tr><td>Two roads diverged in a yellow wood</td><td>Two roads diverged in a yellow woods and his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the w</td></tr>
<tr><td>And on the pedestal these words appear</td><td>And on the pedestal these words appear.
In dangerous heart of my son
Fain and the stone is the prince is death
For the world with his sons of man
And stood he stood he stood he stood
The stone is the true the world dissuades
For the world with his sons of man
And stood he stood he stood he stood
The stone is the true the world dissuades
For the world with his sons of man
And stood he stood he stood he stood
The stone is the true the world dissuades
For the world with his sons of man
And stood he stood he stood he stood</td></tr>
<tr><td>Shall I compare thee to a summer's day?</td><td>Shall I compare thee to a summer's day?
A mother than the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with his sons.
The hardly the world with hi</td></tr>
</table>

Character-based GRU: Two recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wooden bed Which shall be indulgent ear the air As our outworn the way it shamed the man Agrees the stern and concealed hill. The sun shone in the stag was he was here? A willing sad for see the spring A green new without chief to the Gods And the stone-roof'd hut at the same. Ah me! I see me come for the cause I seek is the stag was he wish'd him he comes. Who to the throne I have the grave; But when the stone condemns the strife The sun shone in the stag was he was here? A willing sad for see the
And on the pedestal these words appear	And on the pedestal these words appears. What startless gone will be stamp Were the stone-roof'd his impious setter-and the general good? The sun shone in the stag was he was here? A willing sad for see the spring A green new without chief to the Gods And the stone-roof'd hut at the same. Ah me! I see me come for the cause I seek is the stag was he wish'd him he comes. Who to the throne I have the grave; But when the stone condemns the strife The sun shone in the stag was he was here? A willing sad for see the spring A green new wit
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Not so! let us two days since in the same. Ah me! I see me come for the cause I seek is the stag was he wish'd him he comes. Who to the throne I have the grave; But when the stone condemns the strife The sun shone in the stag was he was here? A willing sad for see the spring A green new without chief to the Gods And the stone-roof'd hut at the same. Ah me! I see me come for the cause I seek is the stag was he wish'd him he comes. Who to the throne I have the grave; But when the stone condemns t

Character-based GRU: Three recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood In the stone is not yet the strong O self-will recklessly indulged The past like half report of her fate. But we see the world dawn an age More fortunate the rest it halls be maintain'd The sceptre--not reigned long in Messenia when he was her speech and day. Into the world which makes he would have strength the sharp cry of her fate. But we see the world dawn an age More fortunate the rest it halls be maintain'd The sceptre--not reigned long in Messenia when he was her speech and day. Into the
And on the pedestal these words appear	And on the pedestal these words appear. Interpret then to the world dispeopled realm Or conquer'd hunter of the throng Back to the torturing thrones Of the high alone!--the breeze Arrests the dinion maidens O my friends Embolden'd by my lenience thine own. Therefore a secret perilous stamed Where the stone is not yet the same; I self-defence and right on the heart speaks the race Caught to his past we hardly strokes Are never fails the rest it run. Certain in this thoughts thou slew'st thy hostile way. All this we had have enjoin'd
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Ah! Carry back the son of Heracles Together who dissuades him but now. Ah me! the world which makes are only That strange was he with an earnest prayer To his past we had have each to hear the world dispeopled realm Or conquer'd hunter of the throng Back to the torturing thrones Of the high alone!--the breeze Arrests the dinion maidens O my friends Embolden'd by my lenience thine own. Therefore a secret perilous stamed Where the stone is not yet the same; I self-defence and right on the heart s

Character-based GRU: 20% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood Of the world who came to see the stone And so thou who at last the hardly friends Of the pour where thy son and strike their profit to the death I come which be the strong The stone and clasm the prince is these prince it like the heart That morn with his grave I have I head I fate down the world with thee and dreads And what the stone--bot we say at last the grey The soul came to his down Then to the throne of the prince His border of the Gods and his strife The stone with the throne of rule I
And on the pedestal these words appear	And on the pedestal these words appears And the death with the death thou wilt from the past And crown'd by the world to the prince Have thee to an army train'd to all the stroke The story with the count of his speaks The hour of the prince is no more and from thy son and hand And hearts they send with all this reach The silence the royal streams to the prince The hand of the world is gone. Thou lives who thou hast not will not lead to rest In the stone with hands and see the prince is not well. Into the stroke the stone-strecks the
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Not we ask in the store. And who to the deep display The brothers with the children father from thy father's train. Ah me! the stone with a starr'd Heracles still with deed And what I am I see the stone-- That we say the glory which he strike them behold! Then to be the stone-chain'd him here. And the deed of the prince is here. And the prince is the hand of peace! Therefore thou hast before thy heart and pray The priviletures not the tomb Hermes to the willing strike The prince in the Gods alo

Character-based GRU: 50% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods Of the lone--but thou the hand and scant And the chase mother than the mountain-toor be now For their free my son and the speaks the strike and the past And the child the father's profined my line See in the child me the strange Sounds and the father for thee and the spream What so me secret profferer heart For the through the fallen and heart have bears Where the father of the father's hand and his friends Of his sounded heart of the broke And the grave the crace and pursued strong the brife
And on the pedestal these words appear	And on the pedestal these words appear. In the story with all the grave. The eyes purtores here and the prayer The prince the stripe and stood Of pastion the sound the chare and humble streams That stouth and the stroke and see the brize And the friends of the past and soul. Thee and the guest-change in the blood? Though so murder of the spreng Had in a death and the strike and the hearts What heart and the should he ask and been The blood and see the strife? When the stand of Heracles and heart has fail heart shall be all thee and
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? And see thee short and the strong A mountain to the world and his some sole down On the throne shall the seen and from thy father stream Who briant and the true the prince and sure. The prince and the gration and the rocks shall be now the shard The peace of the many of the string; And the stroke that great thou hast thou sad and she the stone His brothers for the father from the strife Of the hath to the world and free Where the father's forman mine. The pend me still the chare and the strine.

Character-based GRU: Bidirectional

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods Of a prompter of the past And stood and watch'd the prince is dead And strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead And strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead And strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead
And on the pedestal these words appear	And on the pedestal these words appears The strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead And strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead And strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Ah! its come to see the prince is dead And strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead And strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead And strange which he the royal side Which the three the prince as the grave Of the past of the past And stood and watch'd the prince is dead

Character-based GRU: 1 Nonlinear Layer

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods And the stone the stream of the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream th
And on the pedestal these words appear	And on the pedestal these words appear The stroke the strength and the stone And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Alas! the strength the same the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the same! And the strike the stream the

Character-based GRU: 2 Nonlinear Layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow woods Of a draw new-restless and but not the sons. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of h
And on the pedestal these words appear	And on the pedestal these words appear. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his fall. The stranger of the stranger of his f

Character-based GRU: All Parameters

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wooded one shown The since had he stand as it should received With the storm! with a sight which the bane And who to the storm'd and sad the stricken compeal'd The stands of the strander of the Gods Of the son and see the shadow'd his side. There to the side of the torrent strife The son and the son as the same. And what he stand and thou hast all make the shadow of the past And the stone of the forests and the same. The strong conspiracy fast of morning on Stand and the son and the soul
And on the pedestal these words appear	And on the pedestal these words appear The shelving one the same of my father's heart The sin and the son and friends Of the thrones and the stream the black of man And the stroke of the design'd. The short to the silence a sigh! A passing there to rule them to the strong Her free them to the sinner the rest And see the shadow'd day. In the stream of the King and say: A perish'd in the sons of men When the son and thou say'st a foes The son and the same of Heaven. So thou not to the world of the morning on Since all the storm thou s
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Not with the story of the father's land So all the sinner to the strong The storm! with his son a last before the son His firmer men all the stone-- And thou say'st a soul and the sons. The since hath in the souls of man As out of the lake-- And the storm! In the sinner the sand Of the strong conscience on the tomb While thou sad at the prince is dead And the son a sight and the sons. One say the strong of the dead his side With a strike the sinner to the gracious side To the son and the son as

Word-based GRU models

Word-based GRU: Baseline

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood Stratford Stratford Stratford Stratford Stratford Stratford Stratford Stratford Stratford his To the Preserver Zeus; let me return And he taught him to his new-made friends. His Lyons courage stands not to enquire And even now my son ah me! my son To make it famous by some Knightly deede. The cloudes of dust which he would apply The Temple in the pure Parnassian gorge._ _Stol'n from Aristophanes._ _Of Robin's reed._
And on the pedestal these words appear	And on the pedestal these words appear mountain-crests Stratford loveliest Hartz The King in yeares preacht to his youthfull sonne The prince snatch'd and his uncles with the swell of his crime-- His blood--to me had he prevail'd his Foe To pytus the king seduced of his crime-- To efface in the city of his crime-- His blood--to me in his stead he died. And the prince single pressing the King The bellowing Drumme the Martch againe doth beat.
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? kingdom radiance Stratford dread arena But thou hast wakte of old age is not yet The world's course proves the world Of the high-pasturing kine-- their breath Of the Lower Hartz among oaks And beechen coverts and copse The leaves are born And on the highest pines; to roam The world which the Gods mock To breake the strong Percullice and the beech

Word-based GRU: Two recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood green surprise. loveliest malign Daphnis and his followers; to preuent The growth of France from the wayes arose The curious Fletcher of his bleeding heart? To Mantineia to take Souldiers to restore The punishment of France that all the race of the family Of men the woody high well-water'd dells There fifteene thousand of their Gentrie dyde The prince in yeares preacht to his youthfull sonne That he was paramount did did did loue;
And on the pedestal these words appear	And on the pedestal these words appear And took the lake. of that hill-town quoth before To find a Warre as he had been a pretty while scholars who had formerly been of his acquaintance. They quickly learning him to the queen of France to try the fire of this truth; and Rustum as the prince Laias comes. this gated You haue not writ to him and his sonne To whom this kingdom I know not where. But thou my mother hadst no more thou Thou hadst no _right_ to see thy friends
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? deed To the Messenian chiefs; the people quoth The mules I saw the stag took gone. There fifteene thousand of their Gentrie dyde To the Stymphalian Lake and wondring it was To find a fortune from the Gods to make. And where his Right with dark-streaming blood; To see them in France with vs and to Such are no word of boast but this I say: A private loss then here founds a nation's peace.

Word-based GRU: Three recurrence layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood strive. wealthy Shakespeare! king-- Two little and France the apostle de liz And took the lake. Two hounds and drown'd. The curious Fletcher and nurse his well-strung Bowe And he taught him how he did
And on the pedestal these words appear	And on the pedestal these words appear The cattle which the Pouldron took the Pile Of the Stymphalian gentian to the water's Of the Stymphalian Lake to settle give the deed. I saw the prince and the litter lords The music of the Mighty Mother will'd to cleaue Messenia and the King across the Gods Fled it had not ever thus His faithful coldly nobly and strong And hurry the strong wit of his death.
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Drink having me let me let them see First shame and France yee to France debarre First Conquer the secret of their neighbours. Where twines to snap; the easy-stretch'd endure. The tail's indomitable rolling And 'neath his chestnut-trees form The inheritor of the _Gentiana Coming swiftly in the jocund Dorians The inheritor of the bow

Word-based GRU: 20% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood world And the slow-ripening of those Drummes to keepe: That he was paramount and that debt's paid; And the good learned friendly the air And traversed that glimmering sheet of life and immortality And all the world hath the Gods secure. And the tired Centaurs come to our And wait to see the Ile of Wight. From Plymmouth next and the Dorian lords Of ivy-plants and the shadows and the beech
And on the pedestal these words appear	And on the pedestal these words appear And I must be heard relate I have been always And I would have it! that I am here? Thou hast forgot with thy Fayerie heart I know thee in thy heart is king! And yet I know not what ground-tone I have not what ground-tone I know but recommence Thou hast forgot to see thy husband Thou hast deserud: my old servant and thou seem'st O pytus my son behold behold I
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Which to perform the King they bountifully lent And in the Battell where he stands.' to him The prince snatch'd from the sacrificer's The man whose to the world which they have found. And we shall be unsatisfied as now; I have been enough alone! But with the offending no better keep In a thousand water-breaks of vapours----again Standest on a couch._ of his foes.

Word-based GRU: 50% Dropout

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood The French which else whose the surer of Heracles The Temple of the King of Cypselus Before they will not spare. to what end? And we shall fly for our eyes and his guard. The pensive stranger's The world is no better way. The world thou art gone A palace of vultures A loss indeed!
And on the pedestal these words appear	And on the pedestal these words appear And the stern white Olympus-peaks and laid him down; And in the valley Arcadian and the Adder's The King in the glens of the world Of the rocks of the Gods Of the rocks of the antler'd Of the moonlit peaks and the caves. Of the rocks of the world. Of the same Port and the same Port And what he taught him to the world
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? A cord or no more. a deed And thou art Homer! thou art come to twine And we shall feel the world which heroes Of his own breast fetter'd And we shall be we are not hurl'd And now my stripling foe and plague And wait the world is smooth the day Of sinners and honour'd with thee and air And we shall be unsatisfied and laid the Gods

Word-based GRU: Bidirectional

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood had The sprightly humor of the Eleusinian king-- And yet the tedious tossing to come And the sea-eagles build of the past; Of noon is broken there by chestnut-boughs And on the hunting-rambles which the world The crackling husk-heaps burn as the strife And on the highest pines; The crackling husk-heaps burn the soul. The white mists rolling like a sea!
And on the pedestal these words appear	And on the pedestal these words appear And he taught him flute-playing. in his widow's ears With questionings about an idle tale And had torn up the imperial name. The peak of Jaman stood. The Gods perhaps allow'd you The Roman noble lay; The guard-watch'd Bear. The guard-watch'd Bear. The guard-watch'd Bear. The guard-watch'd Bear. The guard-watch'd Bear.
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? thy harp And drink'st untroubled slumber to an unnatural deed The Gods command not where the heart speaks clear. I saw the mules and litter in the court Of citizens in holiday attire Women and his ill-starr'd brethren fall! The nursling of the Mighty Mother died And where the soaking springs abound And the straight ashes grow at The inheritor of the dells? Of the appeasing gracious harmony Droops the green agnus-castus In the rippled waters grey Of the lentisk and ilex In the rippled waters grey Of the lentisk and ilex In the rippled waters grey Of the lentisk and ilex

Word-based GRU: 1 Nonlinear Layer

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood in the The Dorians lying to the throne be the way of the Gods. PAUSANIAS _a Physician_. CALLICLES _a young Harp-player_. _Morning. A Pass in the forest region of Etna._ _Alone resting the King of the King design'd. The noon of the Mighty Mother will'd The world does is colder yet! to the Gods And the stanch Furies' never-silent scourge. And the stanch Furies' never-silent scourge.
And on the pedestal these words appear	And on the pedestal these words appear The world does to help of his murderer. The mules of the Mighty Mother will'd The man is desperate that fiery blast And shook the world to the strife Of the sun-loving gentian in the heat Of the high-pasturing kine-- And the feast of the world beloud Of the sun-loving gentian in the heat Of the high-pasturing kine--
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? Vpon a little loves that I rehearse And champion to the throne Laias Laias The world does is colder yet! And thou art gone to stray Down the uncompanion'd way! Of the guest-chamber door!-- Ah! in their meadows lone; The leaves of the dells? The white mountain-birds; the Gods

Word-based GRU: 2 Nonlinear Layers

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood white white single single single single single single single single single single single single single single single righteous righteous righteous righteous search fancy chief small righteous small Mother white But I When Thou I A lips small Mother small Mother green In Thou I Thou I Thou I know Thou I know Thou small small search small small search small search small son search small Mother I son son In But I Thou I Thou I Thou I Thou I know I son son son In But I Thou I Thou Thou Thou Thou Thou Thou Thou Thou art Thou hast To offer have I know not to be administered To this my son thou hast be well? To this my son! to be administered but to thee To murder's and helpless my son to knowe To see this kingdom for my son to knowe To be be in my child thou art arrived. To this my son! to be administered to thee To see this kingdom for my son and carried For I have not denied a thing as I came.
And on the pedestal these words appear	And on the pedestal these words appear they catch The King and that countenance had not me. The mules who have lived in vogue The world in their souls to obey The world in the world which we The world in their souls The world in their souls to live! The world in their blankets The world in their blankets The world in their blankets
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? I But I have lived in my son and saw! For I have not what I know not themselves. The prince and the world which they will The Temple in the world deride-- And now the world is mined The world in their souls to obey The world in their souls The world in their souls to obey is gone. A new-made world in Thorney Isle

Word-based GRU: All Parameters

Seed	Output
Two roads diverged in a yellow wood	Two roads diverged in a yellow wood The King of France was euer his youthfull might The widow of the King of his death. The prince snatch'd the prince of his lords; And the prince round--when of the square The King of the Eleusinian king-- And the world loves not the one; And the sea-eagles build the grave The nursling of the Eleusinian king-- And took the world subdue.
And on the pedestal these words appear	And on the pedestal these words appear The nursling of the Mighty of his birth The prickers shouted the prince of the race And the same Port the same moment_ ARCAS The comming of the Eleusinian king-- And the stern of mortal wars And the high mountain-pastures of the stream And the world tempts the sun Of the sun of the Hartz Of the Phrygian flutes
Shall I compare thee to a summer's day?	Shall I compare thee to a summer's day? That radiant at Armes with their owne short years And wait the world subdue. The world's course proves the terms Of the Phrygian highlands came; And the world loves not tame; And yet the world to fix our own And wait the world subdue. And the world tempts the world of Heracles And the world lies the world